

The international journal of science / 2 March 2023

nature

ROOT AND BRANCH

Inventory reveals carbon storage in 9.9 billion trees across African dryland

Politics of hate
Can science help to heal increasingly fractious political divisions?

Sea change
Plotting the course towards autonomous ships and ports

Ice age travel
Genomes map out movement of Europe's hunter-gatherers

VISIT…

LANZAROTE Caliente.com

Nature.2023.03.04

[Sat, 04 Mar 2023]

- **This Week**
- **News in Focus**
- **Opinion**
- **Work**
- **Research**
- **Amendments & Corrections**

This Week

- **Syria after the earthquakes: what researchers can do to help** [28 Feb 2023]
  • Equipment and expert aid are urgently needed for 4.7 million people in the country's neglected northwest.
- **Will the world ever see another IPCC-style body?** [01 Mar 2023]
  • Many have sought to copy the IPCC. A new book explains why the panel's all-encompassing scientific assessments are hard to replicate.
- **Ukrainian science has survived against the odds — now let's rebuild together** [20 Feb 2023]
  • Redoubled international support is needed to renew science in my country, both to help those who stayed and encourage those who have fled to return.

- EDITORIAL
- 28 February 2023

Syria after the earthquakes: what researchers can do to help

Equipment and expert aid are urgently needed for 4.7 million people in the country's neglected northwest.



Ali al-Bash, his mother Amina Raslan and their family in a building in Aleppo damaged by the 6 February earthquake.Credit: Louai Beshara/AFP/Getty

It's been three weeks since southern Turkey and northwest Syria were devastated by earthquakes, starting on 6 February. More than 50,000 people

have died, 70,000 are injured and at least 160,000 buildings have totally or partially collapsed.

United Nations organizations and other aid agencies have joined thousands of volunteers in both countries. The World Food Programme is delivering millions of hot meals. The World Health Organization (WHO) is supplying hospitals with painkillers and antibiotics. The World Bank has pledged nearly US$1.8 billion for recovery and reconstruction, and UN secretary-general António Guterres is appealing for $1 billion for Turkey and close to $400 million for Syria, to be distributed to organizations providing food, shelter and education. Separately, the Red Cross/Red Crescent has launched an appeal to raise around $700 million.



Turkey–Syria earthquake: what scientists know

Turkey's government has pledged to rebuild homes within a year. But there's no such help (and much less other assistance) across the border in northwest Syria, where the earthquake killed more than 4,500 people, injured 8,500, demolished around 10,000 buildings and left some 11,000 people homeless. Northwest Syria's inhabitants have no unified government. They're trapped in a war, a legacy of a 2011 revolution — part of the Arab Spring — that was suppressed by the government of President Bashar al-Assad in a military operation backed by Russia. The region is also effectively isolated;

there are only three (temporary) crossing points for supplies to reach people along Syria's 900-kilometre border with Turkey.

Even before the earthquakes, the UN Office for the Coordination of Humanitarian Affairs estimated that more than half of northwest Syria's 4.7 million people are internally displaced, and around 70% of those are living in temporary accommodation (often in tented cities). More than three million do not have enough to eat and one-third are disabled, many because of the war. Health-care facilities have been targeted, leaving only 66 hospitals still functioning, but poorly equipped, amid an ongoing cholera outbreak.

Nature has spoken to researchers, medical personnel and engineers in the region and internationally. They tell us that northwest Syria urgently needs expertise and specialized equipment, in addition to funding. For example, volunteer engineers are going house-to-house tapping walls with hammers to gauge whether a building is inhabitable.The WHO estimates that between them, the region's hospitals have 64 X-ray machines and 73 renal dialysis machines, 7 computerized tomography (CT) scanners and one magnetic resonance imaging (MRI) machine. But researchers have told *Nature* that access to spare parts and expert technical support is as important as the machinery itself to keep things functioning.

There are research projects studying the situation. One is the UK-funded Research for Health Systems Strengthening in Syria, which is investigating how health care is being affected by war, and determining what more-sustainable models of governance and financing could look like, says co-investigator Abdulkarim Ekzayez, an epidemiologist at King's College London.



One MRI for 4.7 million people: the battle to treat Syria's quake survivors

Members of the international research community can also join broader efforts to alleviate the situation. They can donate funds — for example, by responding to international appeals or giving to the organizations listed at go.nature.com/3kq2rpn. They can urge their lawmakers to call for crossing points into northwest Syria to remain open long after emergency humanitarian assistance ends — two of the three are due to close in May, the third in July. If that happens, the region's people will be cut off.

Researchers can urge the WHO to prioritize the region's health needs. The agency is trusted and, through director-general Tedros Adhanom Ghebreyesus, it could do more to use that status to urge all the power brokers on the ground to allow it to work with other experts to deliver both immediate and longer-term help, including much-needed assistance to rebuild homes and health systems.

This tragedy has opened a rare window to provide more international support for people who have been neglected for much too long. Researchers can help to keep that window open.

Nature **615**, 7 (2023)

doi: https://doi.org/10.1038/d41586-023-00571-7

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00571-7

- EDITORIAL
- 01 March 2023

Will the world ever see another IPCC-style body?

Many have sought to copy the IPCC. A new book explains why the panel's all-encompassing scientific assessments are hard to replicate.



Organizations working to solve global challenges have struggled to replicate the IPCC model for comprehensive scientific assessments.Credit: Steffi Loos/Reuters

When it comes to getting decision makers to pay attention to scientific evidence, there are few better — or perhaps better-known — examples than the Intergovernmental Panel on Climate Change (IPCC). Its summary

reports on climate science, impacts and adaptations are read by a wide range of people, in bodies as disparate as companies and campaign groups, as well as, of course, their primary audience: decision makers. Produced every six or seven years, the main IPCC studies have an extra-ordinary reach, informing everything from global climate agreements, such as that negotiated in Paris in 2015, to the school climate-strikes movement Fridays for Future.

This is no small achievement for what is, at its core, a network of hundreds of researchers working in the early or late hours for no payment. They read and summarize thousands of research papers to answer questions such as how much the planet has warmed; what the future projections are for Earth's climate; what the impacts of warming are; how to mitigate climate change; and how the world can better prepare for a warmer future.



IPCC's starkest message yet: extreme steps needed to avert climate disaster

A similar but less well-known IPCC-style network of researchers is the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES). Established in 2012, its studies helped to underpin the Kunming–Montreal Global Biodiversity Framework, which aims to stem our destruction of nature and was signed in December last year. But when it comes to other great global challenges, which are embodied by the United Nations Sustainable Development Goals — inequality, say, or water or food security — there's no advisory research body on a similar scale or with

comparable impact, and none on the immediate horizon. And this is not for want of trying. So why has the model proved so difficult to replicate?

Product of its time

A new book that takes an in-depth look at the IPCC helps to explain why the climate panel and IPBES remain two of a kind — and why we might not see their like again. *A Critical Assessment of the Intergovernmental Panel on Climate Change*, published in December 2022 and edited by political scientist Kari De Pryck at the University of Geneva in Switzerland and human geographer Mike Hulme at the University of Cambridge, UK, was written collaboratively with 33 other social scientists. It includes chapters on the IPCC's formation and governance, who participates in it and what its future might be: an IPCC-style assessment of the IPCC itself. The authors describe an organization that was a product of a particular set of circumstances, some of which cannot be recreated easily — if at all.

Although the IPCC was established in 1988, the seeds had been sown several decades earlier by researchers who were becoming concerned about the implications of putting vast amounts of carbon into the atmosphere. A series of climate research programmes and conferences were co-organized by the World Meteorological Organization, the International Council of Scientific Unions and the UN. By the late 1970s, studies were starting to accumulate describing projected warming due to atmospheric levels of carbon dioxide and other greenhouse gases.



[Why climate finance is a political hot potato and what to do about it]

Those who would eventually establish the IPCC realized that the findings of such studies could not just be one-offs, and that periodic assessments were needed to keep a closer eye on how the climate was changing. These studies needed to include the science of greenhouse-gas emissions; the nature and composition of greenhouse gases; and the all-important question of the extent to which human activities (particularly the burning of fossil fuels) were responsible for climate change. The last of these was not settled until the mid-1990s, and was reported in the IPCC's second assessment report in 1995, with the now-famous phrase "the balance of evidence suggests that there is a discernible human influence on global climate".

The novelty of the IPCC approach is that researchers employed by participating governments are among the members of the teams reviewing the literature. Scientists representing governments sign off the 'summary for policy-makers', which synthesizes the research into a booklet using language that can be understood by non-experts.

The IPCC's founding researchers considered government involvement to be essential to both attracting decision makers' attention to the body's findings and gaining their trust. But there are good reasons that other research networks cannot replicate this level of involvement.

Scientists working for governments were among the founders of the IPCC because they were already involved in climate science. These were the employees of official weather-data centres, or meteorological offices. They were among the first researchers to have access to the kind of computing power needed for climate simulation studies — which governments tended to have. This meant that it wasn't much of a stretch to include more of them on assessment teams working with researchers at universities.

Many of these government scientists also had strong links to departments for defence — for which accurate weather forecasting is a must — and, through that, had access to some of the most senior people in government. In the United Kingdom, for example, the IPCC's first assessment report (in 1990) was presented in a special seminar to then-prime minister Margaret Thatcher.

Other scientific networks do not have the same degree of access. Even if they did, such involvement would be more complicated to navigate now than in the 1980s. The book's authors detail how, over time, IPCC meetings became more politicized as government representatives — mainly, but not exclusively, from oil-producing states — interfered in the scientists' discussions.

The IPCC's creation was an inventive response by researchers to the challenge of communicating climate science to decision makers. It has brought knowledge of the latest climate science to millions of people. Its reports confirmed a human fingerprint in global warming and, in that regard, have been pivotal to the creation of international climate agreements. From early on, scientists from low- and middle-income countries were invited to join the leadership of the IPCC's writing groups, an unusual move for its time. Nonetheless, the IPCC has struggled to properly represent female researchers, scientists from lower- and middle-income countries, and Indigenous people.

Today's myriad global challenges need research-based evidence more than ever, and it's clear that faster, more focused and more inclusive assessments — such as those that the IPCC and IPBES are starting to do — are the way forward. The world might never again see a research assessment on the scale

that the IPCC pioneered, but that is far from the only way for policymakers to access — and act on — scientific evidence.

Nature **615**, 7-8 (2023)

doi: https://doi.org/10.1038/d41586-023-00572-6

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00572-6

- WORLD VIEW
- 20 February 2023

Ukrainian science has survived against the odds — now let's rebuild together



Redoubled international support is needed to renew science in my country, both to help those who stayed and encourage those who have fled to return.

- Svitlana Arbuzova

For my colleagues and me, the war in Ukraine did not begin on 24 February 2022. It started in 2014, when Russia installed a puppet government in the Donbas in the east of my country. The Eastern-Ukrainian Center for Medical Genetics and Prenatal Diagnosis, a state-of-the-art facility which I founded and have headed for many years, was originally located in Donetsk, the capital of the Donbas.

My colleagues and I publicly expressed our pro-Ukrainian stance and refused to become an institution of the so-called Donetsk People's Republic. We moved to Mariupol, 100 or so kilometres south in unoccupied Ukraine, and restarted work.

In March 2022, soon after the full Russian invasion, our new home was completely destroyed, as was the building in which my colleagues and I rented apartments. Some of us managed to escape the city within days. Others, in particular those unable to leave elderly parents who had moved with them, stayed back in their basements waiting for help. We faced weeks of uncertainty as to their fate.



Ukraine's scientists need help to rebuild their research system

Once again, we survived all the adversities. Our centre might have gone, but our knowledge and experience wasn't captured. Some colleagues with children moved abroad, and others settled in friendly institutions in western Ukraine and continue to collaborate remotely. Our centre now operates in Kyiv, conducting prenatal screening and other types of genetic testing.

When the full-scale war started, it was the first and only moment in my life when I regretted I was no longer young and could not be among those scientists who joined the army fighting for fundamental values most readers of this journal take for granted: freedom and independence. I donated my

car to volunteers, among whom there are many scientists, in particular my close friend physicist Anton Senenko.

I felt relief every time I read his updates of people who could be rescued, from Bucha just outside Kyiv, site of a notorious massacre in the early stages of the war, and from other places. Some scientists we knew have paid the ultimate price: just two weeks ago, it was confirmed that Bizhan Sharopov, a brilliant young physiologist colleague who had been missing in action since April, had been killed.

The courage that Ukrainian people are demonstrating is beyond words, and not only those on the battlefields. Suffering the grief of losing family members, staying in basements without water and food, living now often without light and heat, the whole Ukrainian people is enduring severe deprivation with dignity and unbroken spirit.



Seven ways the war in Ukraine is changing global science

Ukraine's scientists, too, continue our research and are integrated as far as possible into the global scientific community. Some of us have received foreign research grants. I was lucky to obtain the Researchers at Risk Fellowship established by the British Academy and the Royal Society with the Council for At-Risk Academics, allowing me to persevere in my work

researching the causes of pre-eclampsia. I hope that the results will serve my country after victory.

The importance of science, technology and innovation to this future is paramount. Ukraine has many excellent researchers whose work is appreciated by the global scientific community. We must receive a clear signal from our politicians that we are needed, and that science will have a proper place in our national reconstruction.

The Scientific Committee of the National Council of Ukraine for Science and Technology Development, of which I am a member, advises the Ukrainian government, and has already developed proposals for the revival of science. The first priority should be a full assessment of the state of educational and scientific infrastructure, and collaboration with international partners to restore what has been destroyed.



[How three Ukrainian scientists are surviving Russia's brutal war]

The devastating consequences of the war are superimposed on pre-existing problems. The Ministry of Education and Science, which is responsible for assessing the quality and integrity of Ukrainian scientific research, is not fit to perform those tasks. The minister of education and science was himself accused of plagiarism in his earlier scientific work, an allegation upheld by the Ukrainian National Agency for Quality Assurance of Higher Education,

although the matter has since been contested on a procedural technicality in the courts. There must be zero tolerance of any corruption that could hinder the development of the country, especially against the background of Ukraine's candidacy for European Union membership.

The international scientific community's support of Ukrainian scientists has been strong. But current programmes, aimed primarily at researchers who fled Ukraine, should be complemented by a greater focus on those who stayed. There are hopeful signs. In February, the European Commission announced that it would open an office in Kyiv for Horizon Europe, the EU's key research-funding programme. The National Research Foundation of Ukraine, launched in 2018, has administered seven calls for research and development proposals, two of them during wartime, and has partnerships with the Swiss National Science Foundation, the Dutch Research Council and the University of Cambridge, UK. Other institutions and universities should now support and add to these initiatives, not only to expand opportunities for scientists who stayed in Ukraine, but also to attract those who are abroad to return home. Слава Україні!

Nature **615**, 9 (2023)

doi: https://doi.org/10.1038/d41586-023-00518-y

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00518-y

| Section menu | Main menu |

News in Focus

- **Third patient free of HIV after receiving virus-resistant cells** [21 Feb 2023]
  • But the risks associated with the procedure mean it is unlikely to be used widely in its current form.
- **Big dino, little dino: how T. rex’s relatives changed their size** [23 Feb 2023]
  • ‘Impressive’ fossil analysis reveals why some dinosaurs were massive but their cousins were tiny.
- **How will AI change mathematics? Rise of chatbots highlights discussion** [17 Feb 2023]
  • Machine-learning tools already help mathematicians to formulate new theories and solve tough problems. But they’re set to shake up the field even more.
- **US lawsuit threatens access to abortion drug: the science behind the case** [23 Feb 2023]
  • Judge’s decision could ban mifepristone across the country, and weaken the Food and Drug Administration’s authority.
- **Was famed poet Pablo Neruda poisoned? Scientists warn case not closed** [22 Feb 2023]
  • Forensic investigation uncovers evidence that a lethal bacterium could have been in his body when he died.
- **Electrodes build themselves inside the bodies of live fish** [23 Feb 2023]
  • Substance that transforms into a conductive polymer using the body’s own chemistry could improve implantable electronics.
- **How Nature readers are using ChatGPT** [20 Feb 2023]
  • Eighty per cent of respondents have used AI chatbots — and 57% say they use them for ‘creative fun’.
- **How did life begin? One key ingredient is coming into view** [28 Feb 2023]
  • A Nobel-prizewinning scientist’s team has taken a big step forward in its quest to reconstruct an early-Earth RNA capable of building proteins.

- **How to tackle political polarization — the researchers trying to bridge divides** [01 Mar 2023]
  • Political divisions are intensifying, threatening democracies around the world. What strategies bring people closer together again?

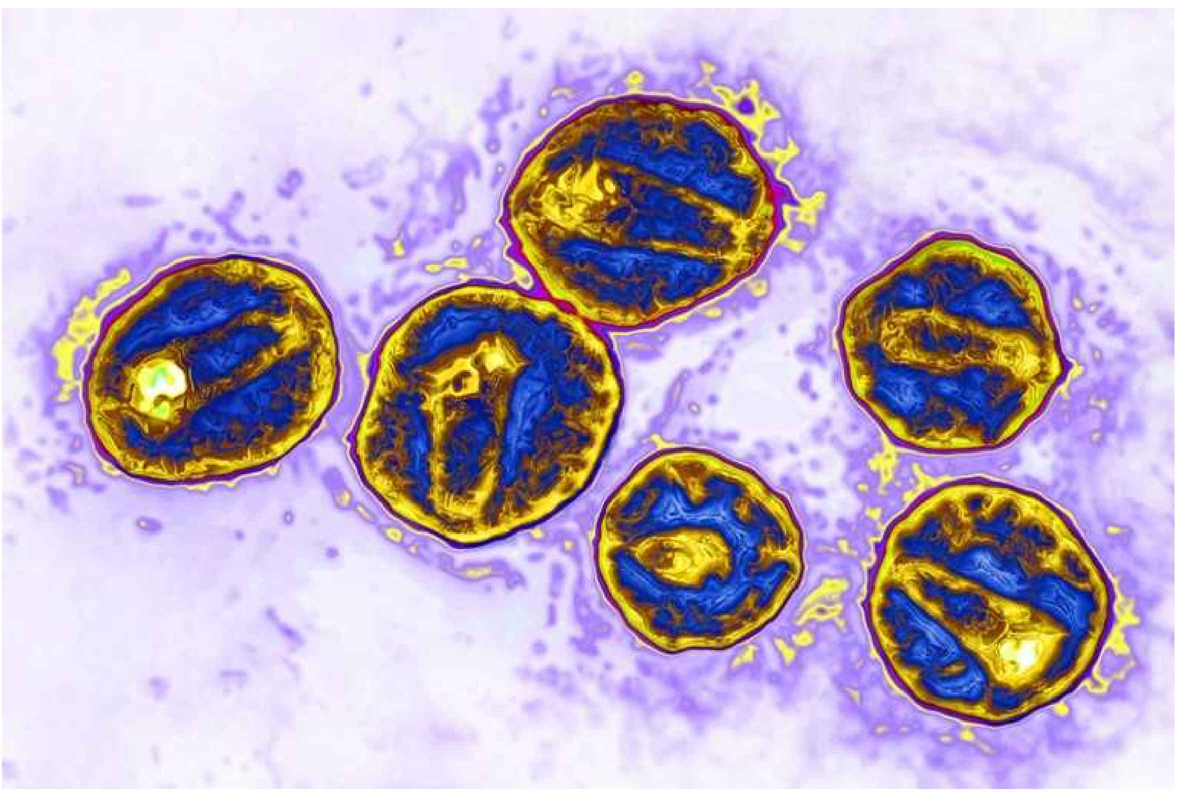
| Next section | Main menu | Previous section |

- NEWS
- 21 February 2023

Third patient free of HIV after receiving virus-resistant cells

But the risks associated with the procedure mean it is unlikely to be used widely in its current form.

- Sara Reardon



HIV particles (pictured) use a receptor called CCR5 to infect human cells. A particular mutation in this receptor makes cells resistant to the virus.Credit: James Cavallini/Science Source/SPL

A 53-year-old man in Germany has become at least the third person with HIV to be declared clear of the virus after a procedure that replaced his bone marrow cells with HIV-resistant stem cells from a donor.

For years, antiretroviral therapy (ART) has been given to people with HIV with the aim of lowering the virus to almost undetectable levels and preventing it from being transmitted to other people. But the immune system keeps the virus locked up in reservoirs in the body, and if an individual stops taking ART, the virus can begin replicating and spreading.

A true cure would eliminate this reservoir, and this is what seems to have happened for the latest patient, whose name has not been released. The man, who is being referred to as the 'Düsseldorf patient', stopped taking ART in 2018 and has remained HIV-free since.

The stem-cell technique involved was first used to treat Timothy Ray Brown, often referred to as the Berlin patient. In 2007, he had a bone marrow transplant, in which those cells were destroyed and replaced with stem cells from a healthy donor, to treat acute myeloid leukaemia. The team treating Brown selected a donor with a genetic mutation called CCR5Δ32/Δ32, which prevents the CCR5 cell-surface protein from being expressed on the cell surface. HIV uses that protein to enter immune cells, so the mutation makes the cells effectively resistant to the virus. After the procedure, Brown was able to stop taking ART and remained HIV-free until his death in 2020.

In 2019, researchers revealed that the same procedure seemed to have cured the London patient, Adam Castillejo. And, in 2022, scientists announced that they thought a New York patient who had remained HIV-free for 14 months might also be cured, although researchers cautioned that it was too early to be certain.

Ravindra Gupta, a microbiologist at the University of Cambridge, UK, who led the team that treated Castillejo, says the latest study "cements the fact that CCR5 is the most tractable target for achieving a cure right now".

Low virus levels

The Düsseldorf patient had extremely low levels of HIV, thanks to ART, when he was diagnosed with acute myeloid leukaemia. In 2013, a team led by virologist Björn-Erik Jensen at Düsseldorf University Hospital in Germany destroyed the patient's cancerous bone marrow cells and replaced them with stem cells from a donor with the CCR5Δ32/Δ32 mutation[1].

Over the next five years, Jensen's team took tissue and blood samples from the patient. In the years after the transplant, the scientists continued to find immune cells that specifically reacted to HIV, which suggested that a reservoir remained somewhere in the man's body. It's not clear, Jensen says, whether these immune cells had targeted active virus particles or a "graveyard" of viral remnants. They also found HIV DNA and RNA in the patient's body, but these never seemed to replicate.

In an effort to understand more about how the transplant worked, the team ran further tests, which included transplanting the patient's immune cells into mice engineered to have human-like immune systems. The virus failed to replicate in the mice, suggesting that it was nonfunctional. The final test was for the patient to stop taking ART. "It shows it's not impossible — it's just very difficult — to remove HIV from the body," Jensen says.

The patient who received the treatment said in a statement that the bone marrow transplant had been a "very rocky road", adding that he planned to devote some of his life to supporting research fundraising.

Timothy Henrich, an infectious-disease researcher at the University of California, San Francisco, says the study is very thorough. That several patients have been successfully treated with a combination of ART and HIV-resistant donor cells makes the chances of achieving an HIV cure in these individuals very high.

Gupta agrees, although he adds that in some cases the virus mutates inside a person and finds other ways to enter their cells. It's also unclear, he says, whether the chemotherapy that the people received for their cancer before their bone marrow transplants might have helped to eliminate HIV by preventing infected cells from dividing.

But it's unlikely that bone marrow replacement will be rolled out to people who don't have leukaemia, because of the high risk associated with the procedure, particularly the chance that an individual will reject a donor's marrow. Several teams are testing the potential to use stem cells taken from a person's own body and then genetically modified to have the CCR5Δ32/Δ32 mutation[2,3], which would eliminate the need for donor cells.

Jensen says that his team has performed transplants for several other people affected by both HIV and cancer, using stem cells from donors with a CCR5Δ32/Δ32 mutation, but that it is too early to say whether those individuals are virus-free. His team plans to study whether, if a person has a larger reservoir of HIV at the time of receiving a transplant, this affects how well the immune system recovers and eliminates any remaining viruses from the body.

Nature **615**, 13-14 (2023)

doi: https://doi.org/10.1038/d41586-023-00479-2

References

1. Jensen, B.-E. O. *et al. Nature Med*. https://doi.org/10.1038/s41591-023-02213-x (2023).

2. Holt, N. *et al. Nature Biotechnol.* **28**, 839–847 (2010).

3. Xu, L. *et al. N. Engl. J. Med.* **381**, 1240–1247 (2019).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00479-2

- NEWS
- 23 February 2023

Big dino, little dino: how *T. rex*'s relatives changed their size

‘Impressive’ fossil analysis reveals why some dinosaurs were massive but their cousins were tiny.

- Dyani Lewis



Theropod dinosaurs such as *Tarbosaurus bataar* grew large or small in a range of ways.Credit: Marco Ansaloni/SPL

A sweeping analysis of shin bones has given researchers a glimpse into how some dinosaurs evolved into mega-beasts such as *Tyrannosaurus*, and others into smaller, bird-like creatures. The work, published this week in *Science*[1], reveals that dinosaurs used more than one evolutionary trick to become larger — or smaller — over time.

Prevailing wisdom held that large-bodied animals are bigger than their smaller-bodied relatives because they grow faster during their most rapid period of growth. That trend holds true for modern animals including birds and mammals — elephants and ostriches grow faster than chihuahuas and sparrows, for example.

It's not the case for all animals. Crocodiles and alligators, for instance, become large because they grow for a long time. But palaeontologists had assumed that for theropod dinosaurs — a group that includes the iconic *T. rex* and which spawned modern birds — large species got big through rapid growth spurts. "It's kind of become the established idea in dinosaurs," says palaeontologist Michael D'Emic at Adelphi University in Garden City, New York.

But that's not what D'Emic found when he sawed into the bones of *Majungasaurus*, a 7-metre-long *T. rex* relative that lived 66 million years ago on what is now Madagascar. The speed of growth in dinosaurs is recorded in rings laid down each year in their bones. Instead of seeing wide rings corresponding to a rapid adolescent growth spurt, D'Emic found lots of narrow growth rings, suggesting that *Majungasaurus* had become large over a prolonged period.

"I was very surprised," he says. The next dinosaur he examined, a similar-sized beast called *Ceratosaurus*, was the opposite — a big dinosaur that grew fast during its growth spurt, says D'Emic.

Bone growth rings

Over a decade, D'Emic and his colleagues amassed bone growth-ring measurements from 42 theropod species to see which strategies led to large and small bodies. They found that 31% of theropod species were larger than

their ancestors because of faster growth and 28% because of prolonged growth. Meanwhile, 21% became smaller than their ancestors by shortening their growth spurts, and 19% by slowing growth.

The study covered theropod species that lived between 230 million years ago and the end of the Cretaceous period 66 million years ago, when a mass-extinction event wiped out the non-avian dinosaurs. It’s “a huge evolutionary timescale”, to include in an analysis, says Vera Weisbecker, an evolutionary biologist at Flinders University in Adelaide, Australia. “That is really impressive,” she says. “It’s just fascinating that there are so many developmental ways to become big or small.”

Palaeontologist Kevin Padian at the University of California, Berkeley, says the analysis is the kind of work that needs to be done, animal group by animal group, to understand how body size evolves.

Drivers of change

But Meike Köhler, an evolutionary palaeobiologist at the Catalan Institution for Research and Advanced Studies in Barcelona, Spain, says the findings are not surprising because previous work has shown a range of growth strategies across animal species. Köhler would like to see an analysis that considers what ecological circumstances influenced how animals changed in size over time.

Weisbecker says that the growth strategy used might be related to evolutionary pressures. “If you looked at all the ones with explosive early growth, you might be able to test if they happen to be the ones that are more likely to be predated on, for example,” she says.

For each species, the growth strategy that led to its individual body size probably related to its unique environment, says Padian. “It’s not a one-size-fits-all, which is a good thing for us to learn,” he says. “We might have thought that, but they’ve documented it.”

D’Emic says he and his team are conducting similar analyses on other groups, including mammals — a group that contains many more species to

sample — to see whether the diversity is found in other branches of the evolutionary tree.

Nature **615**, 14-15 (2023)

doi: https://doi.org/10.1038/d41586-023-00549-5

References

1. D'Emic, M. D. *et al. Science* **379**, 811–814 (2023).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00549-5

- NEWS
- 17 February 2023

How will AI change mathematics? Rise of chatbots highlights discussion

Machine-learning tools already help mathematicians to formulate new theories and solve tough problems. But they’re set to shake up the field even more.

- Davide Castelvecchi



AI tools have allowed researchers to solve complex mathematical problems.Credit: Fadel Senna/AFP/Getty

As interest in chatbots spreads like wildfire, mathematicians are beginning to explore how artificial intelligence (AI) could help them to do their work. Whether it's assisting with verifying human-written work or suggesting new ways to solve difficult problems, automation is beginning to change the field in ways that go beyond mere calculation, researchers say.



DeepMind AI invents faster algorithms to solve tough maths puzzles

"We're looking at a very specific question: will machines change math?" says Andrew Granville, a number theorist at the University of Montreal in Canada. A workshop at the University of California, Los Angeles (UCLA), this week explored this question, aiming to build bridges between mathematicians and computer scientists. "Most mathematicians are completely unaware of these opportunities," says one of the event's organizers, Marijn Heule, a computer scientist at Carnegie Mellon University in Pittsburgh, Pennsylvania.

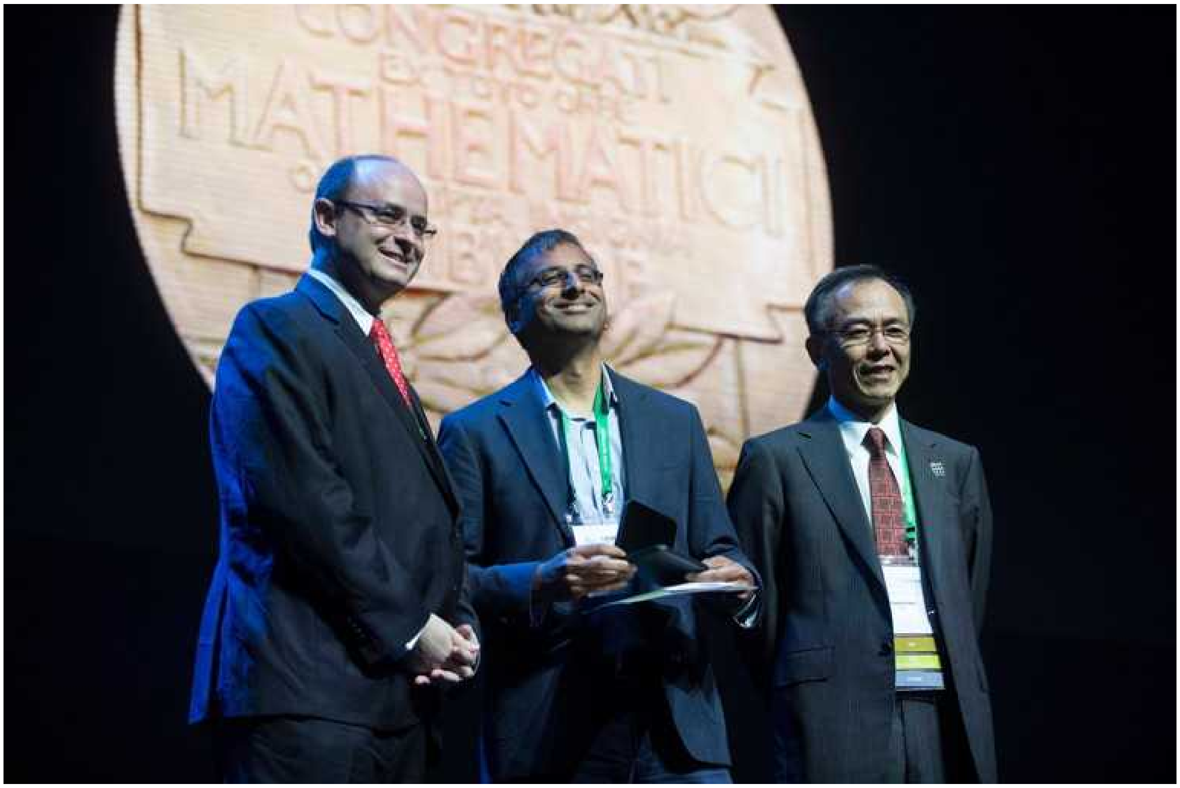
Akshay Venkatesh, a 2018 winner of the prestigious Fields Medal who is at the Institute for Advanced Study in Princeton, New Jersey, kick-started a conversation on how computers will change maths at a symposium in his honour in October. Two other recipients of the medal, Timothy Gowers at the Collège de France in Paris and Terence Tao at UCLA, have also taken leading roles in the debate.

“The fact that we have people like Fields medallists and other very famous big-shot mathematicians interested in the area now is an indication that it’s ‘hot’ in a way that it didn’t used to be,” says Kevin Buzzard, a mathematician at Imperial College London.

AI approaches

Part of the discussion concerns what kind of automation tools will be most useful. AI comes in two major flavours. In ‘symbolic’ AI, programmers embed rules of logic or calculation into their code. “It’s what people would call ‘good old-fashioned AI’,” says Leonardo de Moura, a computer scientist at Microsoft Research in Redmond, Washington.

The other approach, which has become extremely successful in the past decade or so, is based on artificial neural networks. In this type of AI, the computer starts more or less from a clean slate and learns patterns by digesting large amounts of data. This is called machine-learning, and it is the basis of ‘large language models’ (including chatbots such as ChatGPT), as well as the systems that can beat human players at complex games or predict how proteins fold. Whereas symbolic AI is inherently rigorous, neural networks can only make statistical guesses, and their operations are often mysterious.



2018 Fields Medal winner Akshay Venkatesh (centre) has spoken about how computers will change mathematics.Credit: Xinhua/Shutterstock

De Moura helped symbolic AI to score some early mathematical successes by creating a system called Lean. This interactive software tool forces researchers to write out each logical step of a problem, down to the most basic details, and ensures that the maths is correct. Two years ago, a team of mathematicians succeeded in translating an important but impenetrable proof — one so complicated that even its author was unsure of it — into Lean, thereby confirming that it was correct.

The researchers say the process helped them to understand the proof, and even to find ways to simplify it. "I think this is even more exciting than checking the correctness," de Moura says. "Even in our wildest dreams, we didn't imagine that."

As well as making solitary work easier, this sort of 'proof assistant' could change how mathematicians work together by eliminating what de Moura calls a "trust bottleneck". "When we are collaborating, I may not trust what

you are doing. But a proof assistant shows your collaborators that they can trust your part of the work."

Sophisticated autocomplete

At the other extreme are chatbot-esque, neural-network-based large language models. At Google in Mountain View, California, former physicist Ethan Dyer and his team have developed a chatbot called Minerva, which specializes in solving maths problems. At heart, Minerva is a very sophisticated version of the autocomplete function on messaging apps: by training on maths papers in the arXiv repository, it has learnt to write down step-by-step solutions to problems in the same way that some apps can predict words and phrases. Unlike Lean, which communicates using something similar to computer code, Minerva takes questions and writes answers in conversational English. "It is an achievement to solve some of these problems automatically," says de Moura.



AI maths whiz creates tough new problems for humans to solve

Minerva shows both the power and the possible limitations of this approach. For example, it can accurately factor integer numbers into primes — numbers that can't be divided evenly into smaller ones. But it starts making mistakes once the numbers exceed a certain size, showing that it has not 'understood' the general procedure.

Still, Minerva's neural network seems to be able to acquire some general techniques, as opposed to just statistical patterns, and the Google team is

trying to understand how it does that. "Ultimately, we'd like a model that you can brainstorm with," Dyer says. He says it could also be useful for non-mathematicians who need to extract information from the specialized literature. Further extensions will expand Minerva's skills by studying textbooks and interfacing with dedicated maths software.

Dyer says the motivation behind the Minerva project was to see how far the machine-learning approach could be pushed; a powerful automated tool to help mathematicians might end up combining symbolic AI techniques with neural networks.

Maths v. machines

In the longer term, will programs remain part of the supporting cast, or will they be able to conduct mathematical research independently? AI might get better at producing correct mathematical statements and proofs, but some researchers worry that most of those would be uninteresting or impossible to understand. At the October symposium, Gowers said that there might be ways of teaching a computer some objective criteria for mathematical relevance, such as whether a small statement can embody many special cases or even form a bridge between different subfields of maths. "In order to get good at proving theorems, computers will have to judge what is interesting and worth proving," he said. If they can do that, the future of humans in the field looks uncertain.

Computer scientist Erika Abraham at RWTH Aachen University in Germany is more sanguine about the future of mathematicians. "An AI system is only as smart as we program it to be," she says. "The intelligence is not in the computer; the intelligence is in the programmer or trainer."

Melanie Mitchell, a computer scientist and cognitive scientist at the Santa Fe Institute in New Mexico, says that mathematicians' jobs will be safe until a major shortcoming of AI is fixed — its inability to extract abstract concepts from concrete information. "While AI systems might be able to prove theorems, it's much harder to come up with interesting mathematical abstractions that give rise to the theorems in the first place."

Nature **615**, 15-16 (2023)

doi: https://doi.org/10.1038/d41586-023-00487-2

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00487-2

- NEWS EXPLAINER
- 23 February 2023

US lawsuit threatens access to abortion drug: the science behind the case

Judge’s decision could ban mifepristone across the country, and weaken the Food and Drug Administration’s authority.

- Mariana Lenharo



The drugs mifepristone and misoprostol are usually used in combination to terminate a pregnancy.Credit: Jakub Kaminski/East News/Shutterstock

A lawsuit in Texas not only has the potential to further restrict abortion access in the United States — but it could also set a dangerous precedent by overturning the approval of a medication by the US Food and Drug Administration (FDA).



[After Roe v. Wade: US researchers warn of what's to come]

Following [the reversal of *Roe* v. *Wade* last year], some US states have banned abortions, driving more pregnant people to seek medication abortions. The lawsuit against the FDA, brought by anti-abortion groups and physicians, seeks to overturn the agency's 2000 approval of the abortion drug mifepristone. The plaintiffs allege that mifepristone, which is used in combination with another drug, misoprostol, is not safe — a claim that is not corroborated by the scientific evidence, say researchers who spoke to *Nature*. Legal specialists think there is a good chance that the judge deciding the case, Matthew Kacsmaryk in the US District Court for the Northern District of Texas, will rule in favour of the plaintiffs. Appointed by former US president Donald Trump, who promised to help overturn *Roe*, Kacsmaryk "has deep ties to the religious right, and he has issued rulings that are based on very, very conservative ideologies", says Amanda Allen, an attorney and director of The Lawyering Project, an organization based in New York City that works to improve abortion access.

The effects of this case might reverberate across the country, further affecting health care for pregnant people. “If the plaintiffs get what they’re asking for, mifepristone will be banned in all states — it doesn’t matter if the state has a law in place that protects access to abortion,” Allen says.

Here, *Nature* explains the evidence in the case, what’s on the line and what abortion options will be available to people in the United States if the FDA loses.

Is mifepristone safe?

All the evidence suggests that the answer is yes, contrary to what the plaintiffs argue. A 2013 systematic review published in the journal *Contraception*, for example, found that failure to terminate a pregnancy occurred in fewer than 5% of pregnant people who had taken mifepristone combined with misoprostol, and only 0.3% of people were hospitalized after the treatment[1]. The World Health Organization (WHO) lists the regimen as safe, and so does the American College of Obstetricians and Gynecologists (ACOG).



‘We are bearing witness’: Health researchers navigate a post-Roe world

The plaintiffs also allege that the FDA made a mistake when it approved the drug in 2000 and its generic version in 2019. They say that the agency approved the drug using an accelerated process that required it to consider pregnancy an 'illness', for which the abortion drug would provide a 'meaningful therapeutic benefit'. "But pregnancy is not an illness," the plaintiffs state in their complaint.

In response to the lawsuit, the FDA's attorneys have written that mifepristone's approval did not involve an accelerated review (the approval process took four years). They also said that the plaintiffs do not provide any concrete example of a patient who might have suffered serious adverse events associated with mifepristone. "That omission is particularly telling, given the more than two decades that mifepristone has been in use," the court filing says.

If the FDA loses, can people seeking abortion still use misoprostol?

Although the combination of mifepristone and misoprostol is the most commonly prescribed regimen for medication abortion in the United States, both the ACOG and the WHO say that using misoprostol by itself is a safe and effective alternative.

"Mifepristone gets so much attention because it was approved as an abortion drug," says Cari Sietstra, a specialist in reproductive health and justice and principal consultant at Innovations in Reproductive Health Access, a non-profit organization based in Emeryville, California. It was designed to block progesterone, a hormone needed for pregnancy to continue. Misoprostol was originally approved for preventing gastric ulcers. But researchers realized that it causes the uterus to contract, bleed and expel any embryo inside.



Activists march outside the courthouse in Amarillo, Texas, where a lawsuit to ban the abortion drug mifepristone will be decided.Credit: Justin Rex/AP/Shutterstock

So it is an option for medication abortion, says Heidi Moseson, an epidemiologist based in Oakland, California, who works at Ibis Reproductive Health, a global research organization that supports abortion rights. “It’s just that clinicians in the United States have less experience providing it, because for so long the default has been the combined regimen,” she says.

Moseson was the first author of a study published last year[2] in *The Lancet Global Health* that looked into the effectiveness and safety of medication abortion with misoprostol alone or in combination with mifepristone. The study, which followed 961 people who self-managed abortions in Argentina and Nigeria, found that 99% of those who used misoprostol alone had a complete abortion without needing surgical intervention. Among those who used the combined regimen, the rate was 94%.

Another reason why misoprostol isn't often administered on its own is that some clinical trials have shown that it has a lower success rate, Moseson says. But that's probably because of differences in study design and local conditions, she adds. In some studies, if after a few days of taking misoprostol the abortion is not complete, participants are offered a surgical abortion. This would then be registered as an unsuccessful use of the drug. In settings where surgical intervention isn't readily available, the success rates tend to be higher, Moseson says.

The other issue is that some of the previous studies that evaluated misoprostol alone did not use the dosing scheme that is currently endorsed by the WHO. "There's a lot of variation both in the results and in the regimens that are used," says Daniel Grossman, director of Advancing New Standards in Reproductive Health, a social-sciences research programme at the University of California San Francisco.

Could misoprostol eventually be banned?

It's possible. In addition to requesting the ban of mifepristone, the Texas lawsuit invokes an 1873 law, the Comstock Act, which is still a part of the US Code. This act made it illegal to send obscene materials or articles intended for "producing abortion" through the mail. The legal team for the plaintiffs hopes to use this to prohibit the distribution of all abortion drugs by mail.



[Why hundreds of scientists are weighing in on a high-stakes US abortion case]

“Their first stop in this lawsuit is to go after mifepristone. But, make no mistake, the plan here is to end access to medication abortion for everyone in this country,” says Allen.

The precedent set by a court overturning an FDA approval by challenging the agency’s decision-making capacity could endanger misoprostol, too — as well as many medications beyond it, including birth control pills, COVID-19 vaccines and more.

“This really threatens the FDA’s authority over the approval process for medications across the board,” says Allen. “It really kind of flies in the face of the idea that the FDA is there because they’re the ones who have the scientific and medical expertise to make these decisions, not courts.”

What happens next?

Kacsmaryk could make a decision later this month. What happens after that depends on the ruling. If he agrees with the plaintiffs, it could trigger an immediate ban on mifepristone, or he could order the FDA to take certain

steps. In the latter case, “nothing would really happen until those steps were taken. So, we’re really in a state of just waiting and seeing,” Allen says.

If such a decision were to be appealed, it would land in the US Court of Appeals for the Fifth Circuit, “one of the most conservative appellate courts in the country”, according to Allen. Assuming that the court upheld this theoretical decision by Kacsmaryk to ban mifepristone, the next option would be for the FDA to take the case to the Supreme Court.

Nature **615**, 16-17 (2023)

doi: https://doi.org/10.1038/d41586-023-00529-9

References

1. Raymond, E. G. *et al. Contraception* **87**, 26–37 (2013).
2. Moseson, H. *et al. Lancet Glob. Health* **10**, E105–E113 (2022).

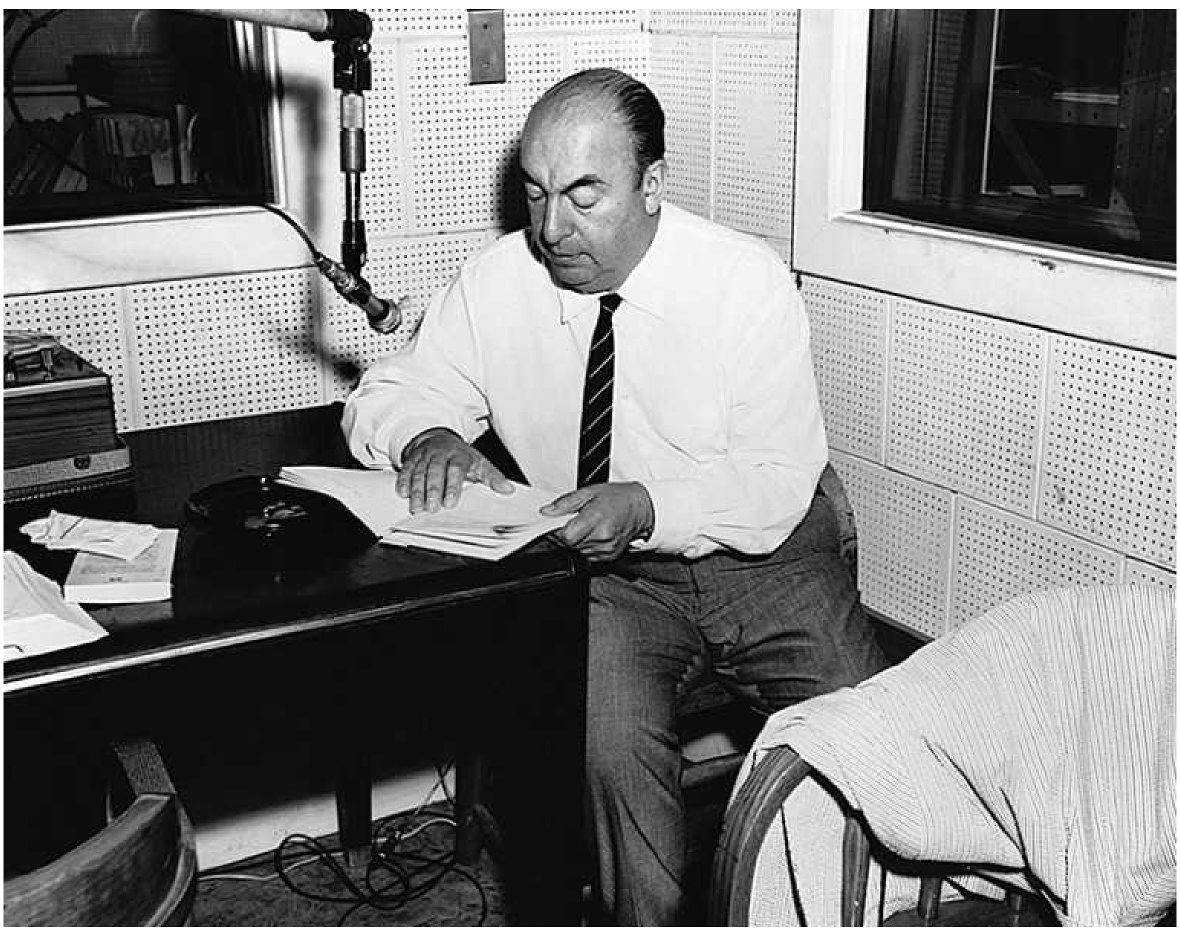
This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00529-9

- NEWS
- 22 February 2023

Was famed poet Pablo Neruda poisoned? Scientists warn case not closed

Forensic investigation uncovers evidence that a lethal bacterium could have been in his body when he died.

- Michele Catanzaro



Pablo Neruda, who won the Nobel Prize in Literature in 1971, reads some of his poetry during a radio interview.Credit: Corbis via Getty

Scientists have concluded that renowned poet Pablo Neruda, a member of the Chilean communist party, might have had a toxic bacterium in his system when he died.

The finding is the latest in a decades-long investigation into the exact cause of Neruda's death on 23 September 1973. Although the poet had advanced prostate cancer when he died, some have said that the timing of his passing — 12 days after general Augusto Pinochet overthrew the socialist government that Neruda supported — was no coincidence.



Can forensics establish whether Pablo Neruda was poisoned?

When announcing the new results on 13 February, Rodolfo Reyes, Neruda's nephew and a lawyer representing his family, said they were proof that his uncle had been poisoned. Remnants of *Clostridium botulinum*, a bacterium that can produce the deadly botulinum toxin, were found in Neruda's teeth, and they could have appeared there as the result of an injection, Reyes said. (Teeth have blood vessels in their roots, so pathogens circulating in a person's blood when they die could theoretically be preserved there.) This method of poisoning "grants the appearance of a natural death", Reyes told *Nature*.

Some of the researchers who helped with the investigation, however, say the evidence is far from conclusive. There is nothing in the science that proves he was poisoned, says Hendrik Poinar, a molecular geneticist at McMaster University in Hamilton, Canada. It's neither a "closed door" nor a "smoking gun", he adds.

A forensic search

In 2011, decades after Neruda's passing, his former driver Manuel Araya alleged in an interview that Neruda had received an unscheduled injection a few hours before dying. Shortly afterwards, the Chilean communist party filed a complaint, saying that Neruda had been poisoned by Pinochet's regime.

To get answers, a judge ordered the exhumation of Neruda's remains and chose an international panel of specialists to analyse them. In 2013, that panel reported that it had screened the remains for some 2,000 chemical agents, such as arsenic, but did not find any evidence of poisoning.



'I feel lost': Chilean researchers saddened by vote to reject new constitution

Two years later, the judge organized a second panel to search for any evidence of biological poisons among the remains. That panel — made up of

researchers from McMaster University, the University of Copenhagen in Denmark, and other institutions — found DNA fragments from *Clostridium botulinum* in Neruda's teeth.

To determine whether the neurotoxin-producing bacterium was in Neruda's body at the time of his death, or whether it was a contaminant that entered from the soil afterwards, the same laboratories continued the investigation. Their work culminated in the latest report, issued to a judge on 15 February. Because the judge has not released the results publicly, *Nature* spoke to some of the scientists who conducted the work.

They say they started by using a technique called shotgun metagenomic sequencing to study human and bacterial DNA extracted from Neruda's bones and teeth. The researchers compared the degradation of the DNA from *C. botulinum* with the degradation of Neruda's own DNA and that of other bacteria in his mouth. They found that the degradation patterns were similar, suggesting that *C. botulinum* could have been in Neruda's body around the time he died.

"Technology has changed dramatically in the last few years. There was no way of doing [all] this seven years ago," Poinar says.

Ruling out the possibilities

Still, the scientists caution against making any strong conclusions from their investigation so far.

"An injection is not the only possible explanation," says Marie-Louise Kampmann, a forensic geneticist at the University of Copenhagen who was part of the panel. For instance, "one alternative would be eating the bacteria in poisoned food", because *C. botulinum* can grow in improperly canned foods, she says.



[Chile's science transformation gains steam with new president]

Kampmann also acknowledges that comparing levels of DNA degradation cannot rule out that the bacterium entered the body after Neruda's death. One DNA sample that has been in a warm, humid environment for only 10 years might have the same amount of degradation as one that's been in a cold, dark setting for 50 years, she says, so comparisons can be deceiving.

Carles Lalueza Fox, an ancient-DNA researcher and director of Barcelona's Natural Sciences Museum in Spain, who was not involved in the studies, agrees. "One cannot completely rule out that the bacteria penetrated in the teeth after death," he says.

The scientists hope the judge will allow more research to be carried out. "We have worked with a heavily degraded and only partial genome," Poinar says. The panel managed to piece together one-third of the bacterial DNA sequence extracted from Neruda's teeth. In that sequence, the researchers identified a gene responsible for producing the botulinum toxin, although they did not find evidence of the toxin itself. "A follow-up is needed to be sure about the presence of [additional] toxic genes" because not all strains of *C. botulinum* produce the toxin, Poinar adds.

It is unclear whether the judge will order further investigations or use the findings to issue a ruling. No deadline has been set for a decision.

Nature **615**, 18-19 (2023)

doi: https://doi.org/10.1038/d41586-023-00525-z

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00525-z

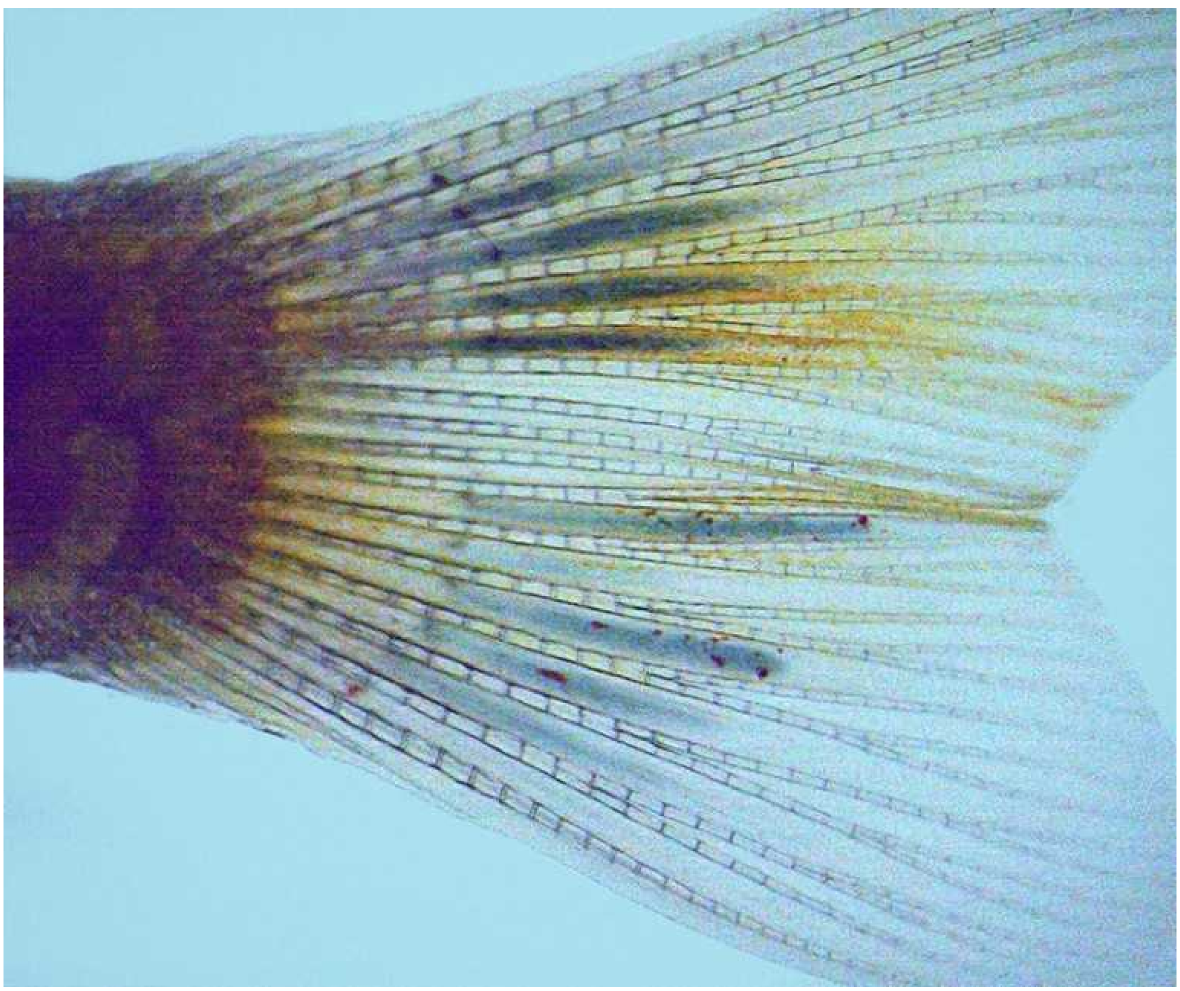
| Section menu | Main menu |

- NEWS
- 23 February 2023

Electrodes build themselves inside the bodies of live fish

Substance that transforms into a conductive polymer using the body's own chemistry could improve implantable electronics.

- Myriam Vidal Valero



Conductive polymers (blue) formed in the tail fins of living zebrafish.Credit: X. Strakosas *et al.*/*Science*

An injectable gel tested in living zebrafish can use the animals' internal chemistry to transform into a conductive polymer.

The discovery, reported on 23 February in *Science*[1], could lead to the development of electronic devices that can be implanted into body tissues such as the brain without causing harm.

When the gel is mixed with the recipient's own metabolites — chemicals generated by the body's processes — a chain reaction turns it into a solid but flexible material.

"We are performing a lot of experiments with these materials to grow electrodes and electronics around cells," says study co-author Magnus Berggren, a materials scientist at the Linköping University in Sweden. He adds that the work could ultimately improve technologies for deep-brain stimulation, for example, or help damaged nerves to regrow.

Internal electronics

Electronic devices or circuitry that can be implanted in the body have many potential applications in medicine and research, such as helping the brain to communicate with prosthetic limbs, or even enhancing memory. But conventional electronic materials can cause inflammation or scarring, and they often deteriorate inside living tissue and eventually stop working.



The brain-reading devices helping paralysed people to move, talk and touch

Although there has been progress towards developing soft, flexible electrodes, it is difficult to get them into the body in a non-invasive way, says Berggren. If you want to insert something deep into the brain, for example, “you will basically make it cut all the way through”, he says.

Berggren’s team wanted to create a material that was conductive, but stable in the long term, non-toxic and of a consistency that allows it to be injected.

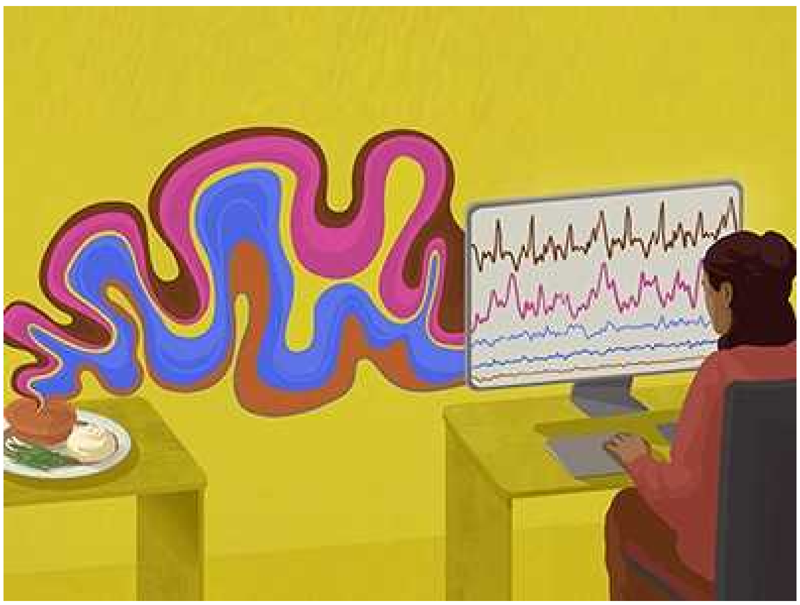
The mixture they developed contains the chemical building blocks for a conductive polymer, along with enzymes. When injected into living tissue, the gel reacts with the common metabolites glucose and lactate, which causes the gel to polymerize into a much firmer — although still soft — material. Working with a group led by chemical biologist Roger Olsson at Lund University in Sweden, the researchers used this approach to generate polymer ‘electrodes’ inside the fins and brains of living zebrafish (*Danio rerio*). They also used it in the nervous tissue of leeches and in muscle tissue from chickens, pigs and cows.

Because the material doesn’t polymerize until it is inside the body, and is “compliant, soft and biocompatible”, it eliminates mechanical differences between typical electrode materials and living tissue that make some medical implants so invasive, says Timir Datta-Chaudhuri, an electrical

engineer at the Feinstein Institutes for Medical Research in Manhasset, New York.

Alternative approach

The idea of using a living tissue's chemistry to create a conducting material inside the body is not new. In 2020, researchers reported engineering an enzyme to be expressed in genetically modified neurons in the nematode worm *Caenorhabditis elegans*. This caused the cells to produce conductive polymers[2].



Restoring smell with an electronic nose

That approach could not be used in people, says Sahika Inal, a bioengineer at the King Abdullah University of Science and Technology in Thuwal, Saudi Arabia. For her, the value of the latest research is that the gel reacts with substances that the body produces naturally, and it does not require the organism to be genetically modified. "I think this technology is providing alternative thinking. Instead of changing the same device's software, why don't we just completely get rid of that device and make the device inside the cell?"

There are still many barriers to be overcome before the injectable substance can be tested in people. Even though the polymer is highly conductive, for example, there is currently no way to make it functional by connecting it to an outside electricity source.

The researchers also need to do more tests to establish that the approach is safe. They didn't observe any unusual behaviour in the zebrafish after injecting the solution into their brains, but they monitored the animals for only three days after the procedure. "They need to look at long-term chronic responses," says Inal.

Nature **615**, 19 (2023)

doi: https://doi.org/10.1038/d41586-023-00544-w

References

1. Strakosas, X. *et al. Science* **379**, 795–802 (2023).
2. Liu, J. *et al. Science* **367**, 1372–1376 (2020).

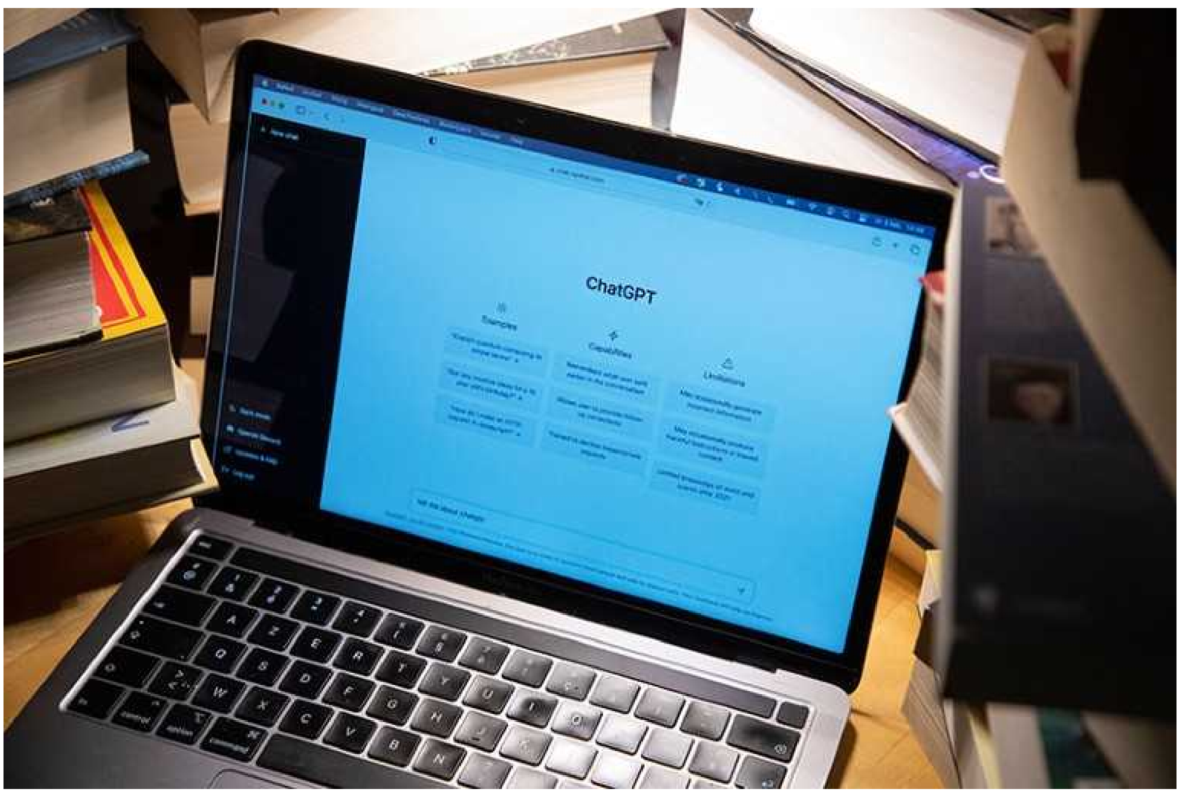
This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00544-w

- NEWS
- 20 February 2023

How *Nature* readers are using ChatGPT

Eighty per cent of respondents have used AI chatbots — and 57% say they use them for ‘creative fun’.

- Brian Owens



The chatbot ChatGPT has been generating debate since it was unveiled in November 2022.Credit: Shutterstock

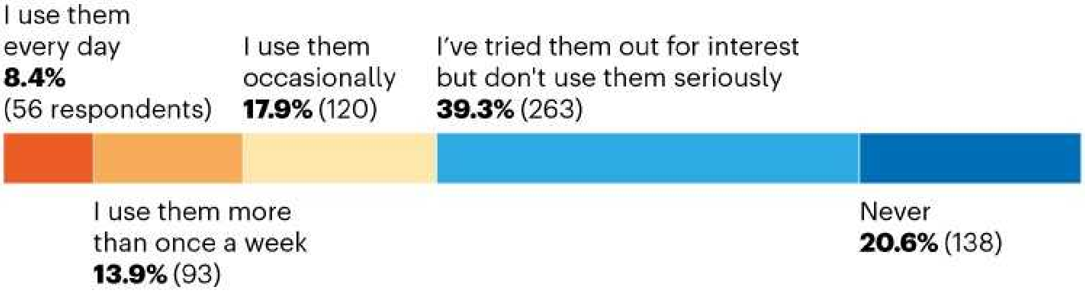
Researchers are keen to experiment with using generative AI tools such as the advanced chatbot ChatGPT to help with their work, according to a survey of *Nature* readers. But they are also concerned about the potential for errors and false information.

Of 672 readers who responded to an online questionnaire, around 80% have used ChatGPT or a similar AI tool at least once. More than one-fifth use such tools regularly — 8% said they use them every day, and 14% several times per week. Around 38% of respondents know of other researchers who use the tools for research or teaching (see 'Chatbot users').

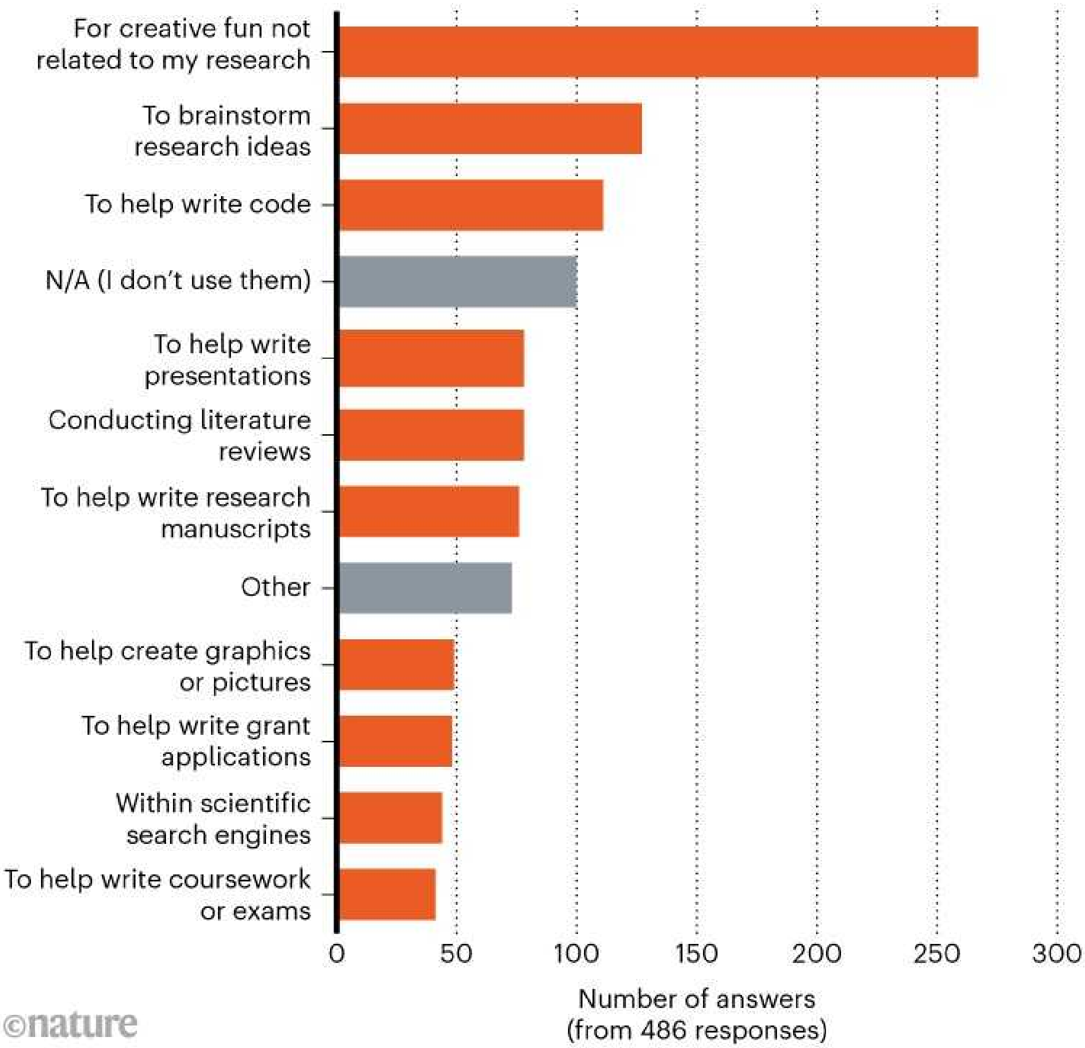
CHATBOT USERS

Most survey respondents have tried generative artificial intelligence (AI) tools, and some said they use them to help with various research-related tasks.

Q: How often do you use ChatGPT, GPT-3, DALL-E, MidJourney, Stable Diffusion, any other 'generative AI' tools, or products built using them?



Q: What do you use these tools for? (Choose all that apply).



©nature

A considerable proportion of respondents — 57% — said they use ChatGPT or similar tools for “creative fun not related to research”. Among the uses related to science, brainstorming research ideas was the most common, with 27% of respondents indicating they had tried it. Almost 24% of respondents said they use generative AI tools for writing computer code, and around 16% each said they use the tools to help write research manuscripts, produce presentations or conduct literature reviews. Just 10% said they use them to help write grant applications and 10% to generate graphics and pictures. (These numbers are based on a subset of around 500 responses; a technical error in the poll initially prevented people from selecting more than one option.)



‘Arms race with automation’: professors fret about AI-generated coursework

Survey participants shared their thoughts on the potential of generative AI, and concerns about its use, through open-ended answers. Some predicted that the tools would have the biggest beneficial impacts on research by helping with tasks that can be boring, onerous or repetitive, such as crunching numbers or analysing large data sets; writing and debugging code; and conducting literature searches. “It’s a good tool to do the basics so you can concentrate on ‘higher thinking’ or customization of the AI-created content,” says Jessica Niewint-Gori, a researcher at INDIRE, the Italian ministry of education’s institute for educational research and innovation in Florence.

Some hoped that AI could speed up and streamline writing tasks, by providing a quick initial framework that could be edited into a more detailed

final version.

"Generative language models are really useful for people like me, for whom English isn't their first language. It helps me write a lot more fluently and quicker than ever before. It's like having a professional language editor by my side while writing a paper," says Dhiliphan Madhav, a biologist at the Central Leather Research Institute in Chennai, India.

Fictitious literature

But the optimism was balanced by concerns about the reliability of the tools, and the possibility of misuse. Many respondents were worried about the potential for errors or bias in the results provided by AI. "ChatGPT once created a completely fictitious literature list for me," says Sanas Mir-Bashiri, a molecular biologist at the Ludwig Maximilian University of Munich in Germany. "None of the publications actually existed. I think it is very misleading."



What ChatGPT and generative AI mean for science

Others were concerned that such tools could be used dishonestly, by cheating on assignments, for example, or by generating plausible-sounding

scientific disinformation. The idea that AI could be used by ‘paper mills’ to produce fake scientific publications was mentioned frequently, as was the possibility that over-reliance on AI for writing tasks could impede researchers’ creativity and stunt the learning process.

The key, many agreed, is to see AI as tool to help with work, rather than to replace work altogether. “AI can be a useful tool. However, it has to remain one of the tools. Its limitations and defects have always to be clearly kept in mind and governed,” says Maria Grazia Lampugnani, a retired biologist from Milan, Italy.

Nature **615**, 20 (2023)

doi: https://doi.org/10.1038/d41586-023-00500-8

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00500-8

- NEWS FEATURE
- 28 February 2023

How did life begin? One key ingredient is coming into view

A Nobel-prizewinning scientist’s team has taken a big step forward in its quest to reconstruct an early-Earth RNA capable of building proteins.

- Amber Dance



Illustration by Hudson Christie

Billions of years ago, before there were beasts, bacteria or any living organism, there were RNAs. These molecules were probably swirling around with amino acids and other rudimentary biomolecules, merging and diverging, on an otherwise lifeless crucible of a planet.

Then, somehow, something special emerged: a simple machine, a pocket made of RNAs, with the ability to place amino acids next to one another and maybe link them into chains. This was the macromolecule that would gradually evolve into the ribosome, the RNA–protein complex responsible for translating genetic information into proteins. Its birth — the details of which remain hypothetical — would have created a fundamental shift in this prebiotic, RNA-dominated world, providing a key ingredient to all life as we know it. Ada Yonath, a structural biologist at the Weizmann Institute of Science in Rehovot, Israel, and her team first conceptualized this 'protoribosome' idea nearly two decades ago, after she and others determined the structure of the modern ribosome, a feat that later secured Yonath a share of the 2009 Nobel Prize in Chemistry.

But to solidify the case for the hypothetical protoribosome, Yonath and her laboratory would have to build it.

It's a project that other scientists have watched with interest. And the lab's achievements in the past two years — creating a primitive RNA machine that can link two amino acids together[1,2] — have created a ripple of excitement. A group in Japan, working separately and led by molecular biologist Koji Tamura at the Tokyo University of Science, has succeeded in creating a similar, functional protoribosome[3]. "It's strongly supportive of Ada's protoribosome idea," says Paul Schimmel, a biochemist at Scripps Research in La Jolla, California. "I've always had an open mind; the open mind is becoming a more convinced mind."

Although doubts and caveats remain, Yonath's and Tamura's work seems to recapitulate a milestone on the road from primordial organic molecules to the ribosome used by the last common ancestor of all living things. This was no simple task: In Yonath's group, the project was passed from researcher to researcher and took more than 15 years to succeed. The work has now opened the door for origin-of-life scientists to fill in further details. And

others are looking at the protoribosome, or something like it, as a tool to create new kinds of biomolecule.

“This should be a starting point of many more fields of research,” says Tanaya Bose, a postdoctoral researcher who has led the Yonath group’s efforts for the past several years.

Crystal inspiration

Scientists have been attempting to recreate some semblance of the chemical origin of biomolecules for decades; it was 70 years ago that Stanley Miller, a chemist at the University of Chicago in Illinois, sparked a gas mixture to create organic compounds[4]. Researchers including Carl Woese and Francis Crick suggested the ribosome could have started as a molecule made solely of RNA, an idea that was supported 40 years ago with the proof that RNAs can catalyse reactions[5,6]. This led to the ‘RNA world’ hypothesis, which describes a time, before cells or actual life, when RNAs replicated and catalysed reactions[7]. The RNA-world idea has come into doubt in the past decade; many scientists now suspect that a variety of biomolecules, including rudimentary proteins, lipids and metabolites, might have existed together with nucleic acids.

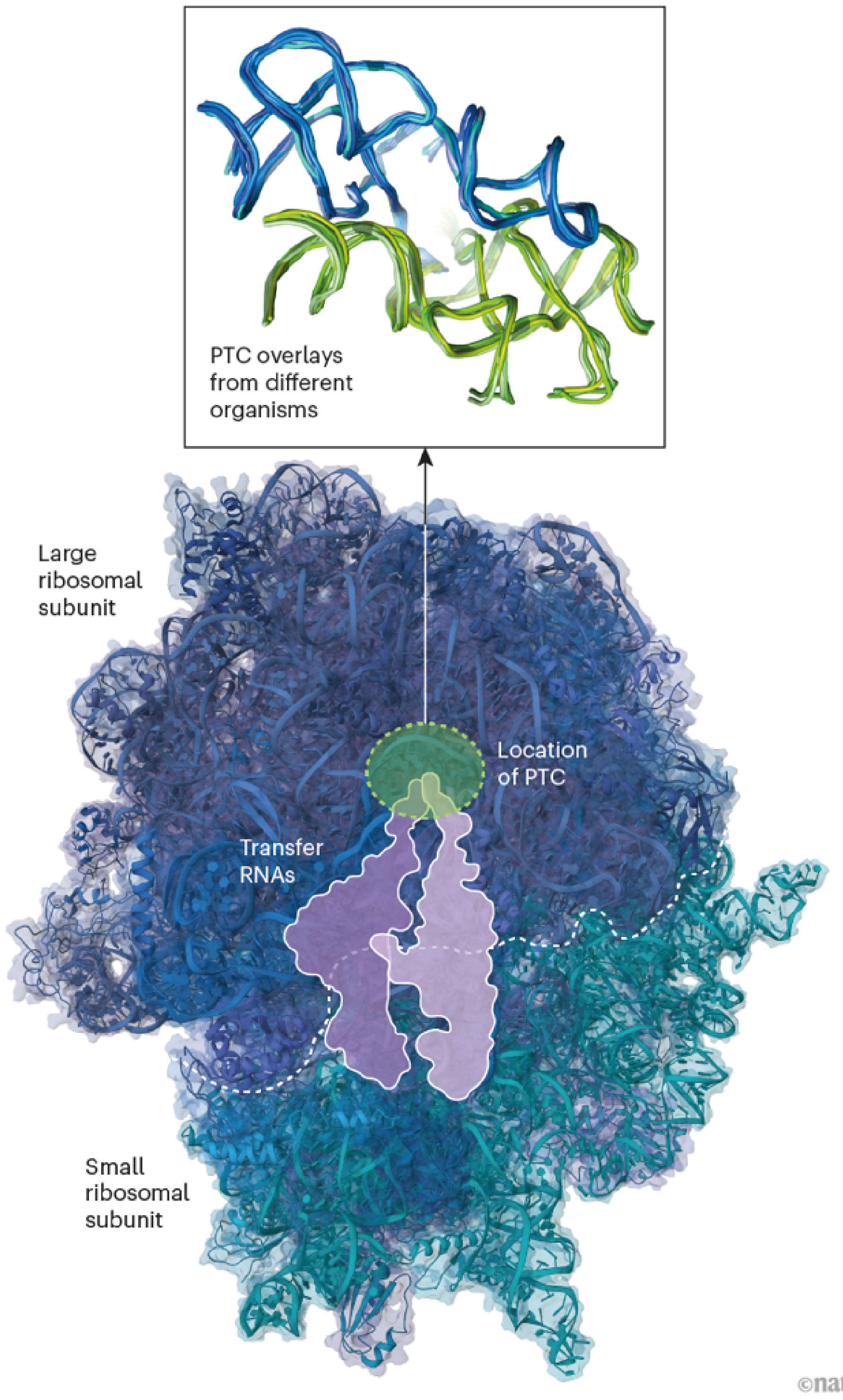
Yonath suspects there were plenty of RNAs on early Earth. “Most of them don’t exist any more because they were small and not useful.” The protoribosome, she says, was the one that lasted.

Yonath and her colleagues published high-resolution structures for the two protein–RNA subunits that make up a ribosome in the early 2000s[8,9]. These examples came from extremophile bacteria. And as ribosomal structures from other organisms were published, Ilana Agmon, a scientific adviser in Yonath’s group, noticed something that she found striking. Deep in the core of the large subunit was a semi-symmetrical segment. This region contained a pocket-like structure, made of ribosomal RNA, called the peptidyl transferase centre (PTC). During the translation of mRNA into protein, when two amino acids are placed in the PTC, it creates conditions for them to click together. And, although the specific nucleotide sequence of the structure varied between species, the shape was the same in every example,

suggesting that it’s crucial to the ribosome’s ability to support life (see ‘The heart of the matter’).

THE HEART OF THE MATTER

At the heart of the ribosome sits the peptidyl transferase centre (PTC), a pocket of RNA in which amino acids, carried by transfer RNAs, come together to form complex proteins. When the 3D structure of the PTC from vastly different organisms — eubacteria, archaea, eukaryotes and even mitochondria — are overlaid, the shape is nearly identical, suggesting it is important to all life.



©nature

Source: refs 1 & 3/PDB.

In 2006, when Yonath started asking her team about the evolution of the ribosome, Agmon suggested looking closely at the semi-symmetrical region containing the PTC. “We speculated that this is the protoribosome, this is the part from which the ribosome evolved,” recalls Anat Bashan, a senior staff scientist who’s been involved in the project since the beginning. But, the region they were positing as the putative protoribosome is assembled from 178 ribonucleotides, which Yonath noted would be an awfully large structure to spring into being, fully formed, on primordial Earth. Based in part on the symmetry she’d observed, Agmon proposed a model requiring 2 similar, L-shaped RNAs of 60 and 61 nucleotides. The team thought this was a more reasonable size of molecule to arise on early Earth.

Fragments of that size are certainly plausible. Elisa Biondi, a biochemist at the Foundation for Applied Molecular Evolution in Alachua, Florida, and her colleagues managed to synthesize RNAs of about 100–300 bases long on materials called rock glasses, which would have formed owing to volcanic action or meteor impacts in a prebiotic world[10].

But not everyone is convinced that Yonath’s and Agmon’s short fragments arose spontaneously. Joseph Moran, an organic chemist at the University of Strasbourg in France, praises Yonath’s accomplishments but doubts that the protoribosome just popped into existence. “It had to be derived from much simpler things,” he says.

Robert Root-Bernstein, a biologist at Michigan State University in East Lansing, has a theory for what those simpler things might have been: transfer RNAs (tRNAs). In the modern ribosome, amino acids enter the PTC attached to tRNAs that match the three-letter codes on the mRNA. Those codes determine which amino acid comes next in the protein. To Root-Bernstein and others, the PTC core looks a lot like four tRNAs joined together. And tRNAs, he notes, don’t just deliver amino acids to the ribosome; they are versatile molecules that can do all kinds of task, such as sensing nutrients and silencing genes. Perhaps they had some function before the protoribosome, then provided the building blocks for that structure.

Extraordinary evidence

Regardless of how the hypothetical protoribosome came to be, at the time when Agmon and Yonath first conceived it, there was no experimental evidence to show that it could have existed and worked as they thought it did. The protoribosome hypothesis assumes that one of those early RNA pockets could link amino acids together, and that it then evolved into the ribosome, which has that same function. Scientists who spoke with *Nature* say that's a fair idea, but it's not a given. "It's plausible," says Anton Petrov, an evolutionary biologist at the Georgia Institute of Technology in Atlanta, but he also thinks that early RNA machines might have had functions distinct from peptide synthesis, and that they then took on that role later as the protoribosome emerged.

"It was a very bold and imaginative suggestion from Yonath," says John Sutherland, a chemist at the Medical Research Council Laboratory of Molecular Biology in Cambridge, UK.



Oldest-ever DNA shows mastodons roamed Greenland 2 million years ago

Bold theories require extraordinary evidence, and that's what the lab set out to obtain. After developing the protoribosome hypothesis, Agmon moved to the Technion — Israel Institute of Technology in Haifa. Graduate student

Chen Davidovich took the lead as the experiments commenced. The first step, Yonath says, was to produce the molecules to build this theoretical protoribosome.

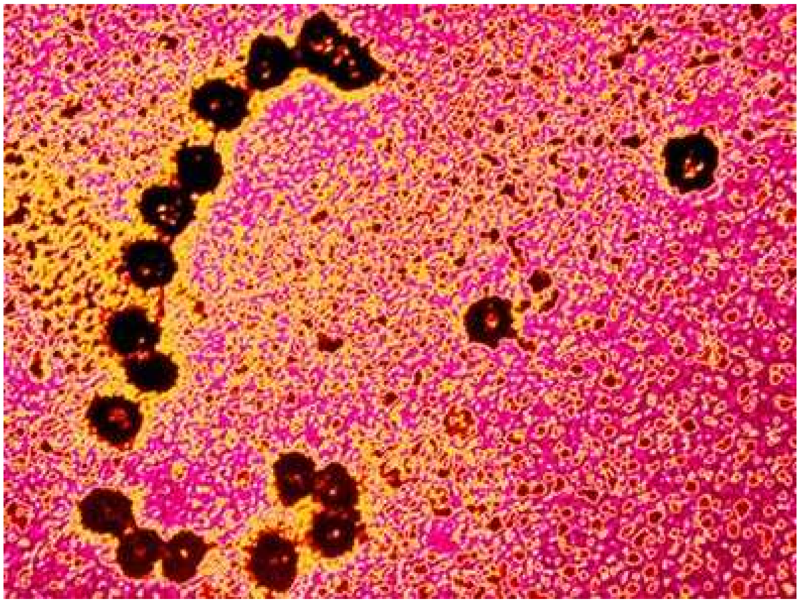
Davidovich studied the RNA sequences of various modern ribosomes. The macromolecule contains dozens of ribonucleotides and accessory proteins, but not all of these are involved in the PTC's shape or functions. He stripped out whatever seemed like it would be extraneous to the protoribosome, leaving just enough RNA to create that semi-symmetrical pocket[2]. Some of these RNAs were able to pair up into something similar to the PTC core that Agmon had imagined.

"This took a long time," Yonath recalls — but the next step would take even longer. The second step was to show that these putative protoribosomes could take two amino acids and hook them together.

Davidovich attempted to show PTC activity in the same way that other researchers have measured the ability of the modern ribosome to link two amino acid analogues into a dipeptide. He labelled such analogues, linked to a handful of nucleotides to stand in for tRNA, with radioactivity. After mixing them with the protoribosome, he thought he would be able to sort the molecules by size and find longer, radio-labelled dipeptides. But he never saw a hint of dipeptide. "It almost killed my career," recalls Davidovich. Fortunately, he had back-up projects related to another interest of Yonath's lab, antibiotics that attack ribosomes. He graduated in 2010 and now leads a lab studying gene repression at Monash University in Melbourne, Australia.

Next up was graduate student Miri Krupkin, who was drawn in by the symmetry of the PTC and its ability to build proteins quicker than any human chemist has managed. Using Davidovich's constructs and others that Krupkin designed herself, she too attempted to detect the radioactive dipeptide product on the basis of on its size. She also tried labelling the substrates with fluorescence, instead of radioactivity. Still, nothing.

Krupkin began to wonder if the protoribosome even made peptides when it first emerged. She tested other potential chemistries, such as adding or removing phosphate groups from other molecules, and still came up dry.



[Origin of life theory involving RNA–protein hybrid gets new support]

“I kept going, because it’s such an interesting question,” says Krupkin. But similar to Davidovich, she had to rely on another project, also on antibiotics, to graduate in 2016. She’s currently investigating HIV’s RNA structure as a postdoc research fellow at Stanford University in California.

Bose, who joined the lab as a postdoc in 2016, would take the project to the finish line. She had trained as a chemist, and that gave her a fresh perspective. She knew that if the primitive protoribosome worked at all, it would probably be inefficient, yielding a minute amount of dipeptide. Rather than separating the reaction products by size, Bose turned to mass spectrometry, which had advanced to become the most sensitive method by this time. It still wasn’t easy — she performed many reactions and control experiments, and tried two forms of mass spectrometry — but at last, she saw a peak representing the expected dipeptide[1,2].

“The amount of product was tiny,” says Bose. She suspects Davidovich and Krupkin might have been making it all along, but there wasn’t enough to detect it using their methods.

Even that tiny amount was “a bit of a feat” to produce, says Sutherland. “They’ve made progress on a really difficult problem.”

Next steps

By the time Krupkin was trying to measure a dipeptide product, the Yonath group wasn't the only one on the trail of the minimal peptide-bonding machine. Tamura, in Japan, was also inspired by the semi-symmetrical core pocket. He recruited a group including master's student Mai Kawabata and undergraduate Kentaro Kawashima to make their own protoribosome-like structure.

Their PTC facsimile, made from two 74-nucleotide RNAs, was similar to Yonath's, but to stand in for modern tRNAs, they used structures called minihelices. These are about half the size of a modern tRNA, and were therefore considerably larger than the tRNA replacements Yonath used. Modern tRNAs are thought to have evolved from minihelices.

Tamura's team eventually achieved its own success, also detecting the dipeptide with mass spectrometry[3]. "Mai and Kentaro were patient and responsive, and finally, we won a big win," says Tamura.

Schimmel, Tamura's former postdoc adviser, says that the minihelices were a key addition. "He carried it to the next level, so to speak, in that he constructed what is believed by many evolutionists to be the primordial tRNA," says Schimmel. Some scientists, he notes, even suspect that minihelices evolved the ability to self-replicate, another key step on the way to living organisms.



The shifting sands of ‘gain-of-function’ research

Tamura cautions that neither lab’s construct works exactly like the modern PTC does. “All of our results are too simple,” he says. “We still have a long way to go before we really understand the evolution of the PTC and the ribosome.”

For Bashan, it’s close enough. “We think this is how the first proteins came about,” she says.

But as other scientists note, there were probably other ways for polypeptides to emerge on early Earth. Linking up amino acids is “dead easy”, says Sutherland. For example, some scientists have proposed that amino acids plus α-hydroxy acids (a group that includes, for example, lactic acid and citric acid), could have formed into polypeptides during cycles of cool, wet and hot, dry conditions on early Earth — with no RNA required[11]. And last year, another team used RNA bases, although not the A, C, G, and U ones used in modern RNA codes, to put peptides together. The researchers proposed this could have happened, without the ribosome, in the RNA world[12]. The catch, says Sutherland: “it bore no resemblance to the way that nature now makes peptides from RNA”.

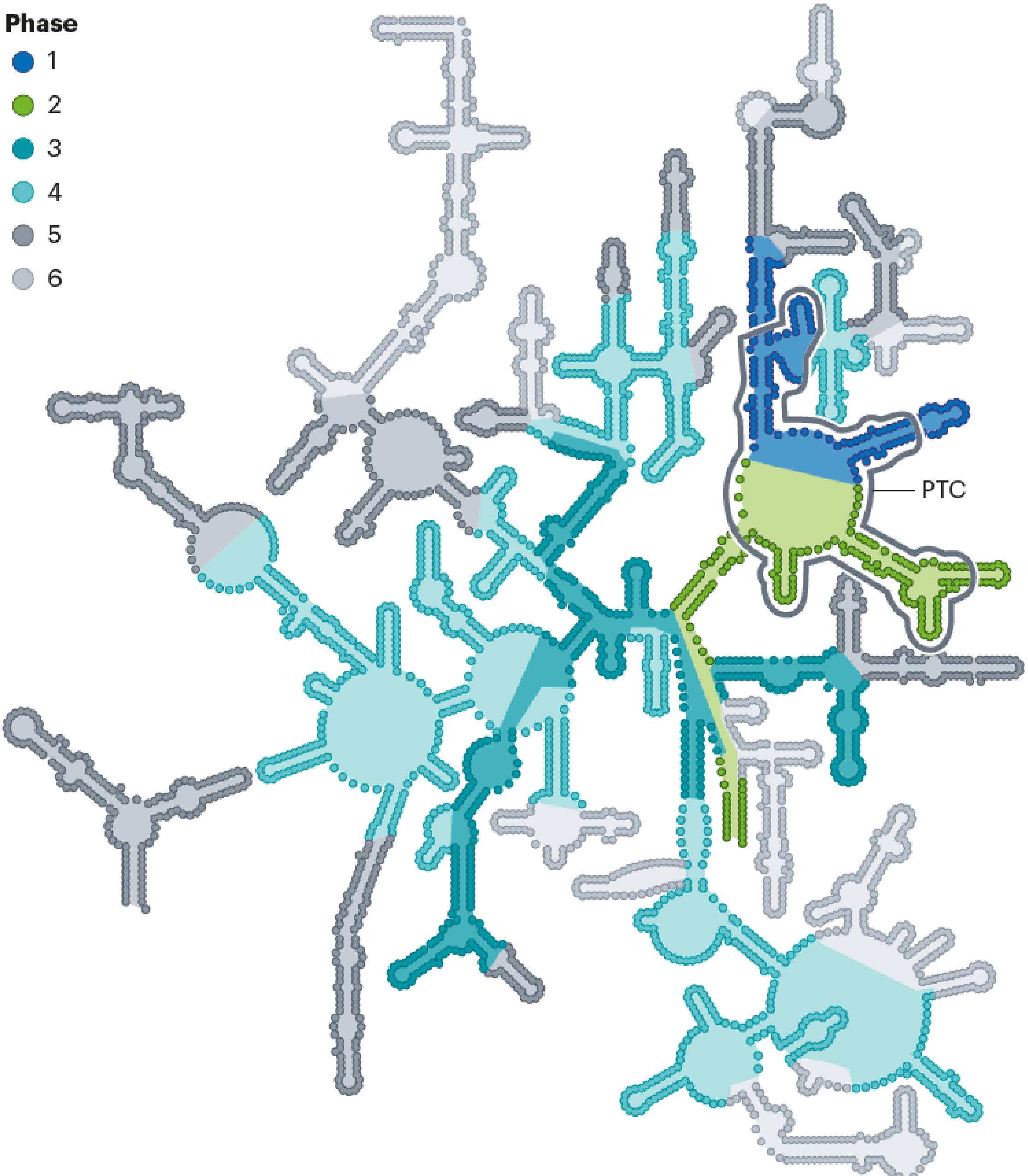
What is special about the protoribosome work, say Sutherland and Yonath, is that it is possible to imagine how this primitive core, over millennia, might

have accumulated extra pieces of RNA and protein to create the modern ribosome.

Petrov and his colleagues have predicted exactly that kind of timeline by working backwards from today's ribosome. They analysed ribosomes for atomic-level 'fingerprints' that were left in the 3D ribosome structure when new branches of RNA were added[13]. This led to a model that showed early structures matching those of Yonath's protoribosome, with extra pieces of RNA slowly incorporated over time (see 'Molecular evolution').

MOLECULAR EVOLUTION

When researchers modelled the evolution of the modern ribosome they found six phases of step-wise expansion in the ribosomal RNA. The earliest phases, coloured dark blue and green here, contain the peptidyl transferase centre (PTC), where amino acids are stitched together.



©nature

Source: ref. 13.

One obvious next step for Yonath's team is to try to make a diverse set of peptides that are longer than two amino acids. Bose is working on this, but it will require different starting materials.

There is one group that claims to have synthesized a longer peptide from a protoribosome-like structure. Yuhong Wang, a biophysicist at the University of Houston in Texas, and her colleague reported mass spectrometry results from experiments with a larger version of the PTC that indicate the production of a chain of nine lysines[14]. Wang thinks that there might be lysine polymers of other lengths as well. "I think we have strong evidence," she says, but she admits that the controls are incomplete, and she cannot explain why strings of nine amino acids would be the most prominent product.

Wang and others are interested in using such stripped-down ribosomes to manufacture new kinds of biomolecule, not limited to the usual 20 amino acids or even to amino acids at all, that could have uses in medicine or industry. For example, they might build molecules out of amino acids that are right-handed in structure, unlike the left-handed ones seen in life on Earth. This kind of macromolecule synthesis could be cheaper and more environmentally friendly than other means, Wang suggests.

Meanwhile, there's plenty of research left to do for origin-of-life studies. Scientists need to work out how RNAs gained the ability to self-replicate. And they need to discover how an early ribosome would have managed to create specific peptides encoded by primitive mRNAs. Those processes, plus the ability to synthesize peptides, would provide the raw materials for evolution.

There's one more factor, Sutherland says: the early peptides made by the protoribosome must have been somehow useful, or there would be no evolutionary advantage to the machine's continued existence. He suggests a couple of speculative functions: perhaps the peptides sequestered metal ions that would otherwise destroy RNAs. Or, they might have helped to form early biomolecular compartments to concentrate RNA and peptides together.

“When you get something that evolution can act on,” says Sutherland, “the rest is history.”

Nature **615**, 22-25 (2023)

doi: https://doi.org/10.1038/d41586-023-00574-4

References

1. Bose, T., Fridkin, G., Bashan, A. & Yonath, A. *Israel J. Chem.* **61**, 863–872 (2021).

2. Bose, T. *et al. Nucleic Acids Res.* **50**, 1815–1828 (2022).

3. Kawabata, M. *et al. Life* **12**, 573 (2022).

4. Miller, S. L. *Science* **117**, 528–529 (1953).

5. Kruger, K. *et al. Cell* **31**, 147–157 (1982).

6. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. *Cell* **35**, 849–857 (1983).

7. Gilbert, W. *Nature* **319**, 618 (1986).

8. Schluenzen, F. *et al. Cell* **102**, 615–623 (2000).

9. Harms, J. *et al. Cell* **107**, 679–688 (2001).

10. Jerome, C. A., Kim, H.-J., Mojzsis, S. J., Benner, S. A. & Biondi, E. *Astrobiology* **22**, 629–636 (2022).

11. Forsythe, J. G. *et al. Angew. Chem. Int. Edn Engl.* **54**, 9871–9875 (2015).

12. Müller, F. *et al. Nature* **605**, 279–284 (2022).

13. Petrov, A. S. *et al. Proc. Natl Acad. Sci. USA* **111**,10251–10256 (2014).

14. Xu, D. & Wang, Y. *Biochem. Biophys. Res. Commun.* **544**, 81–85 (2021).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00574-4

- NEWS FEATURE
- 01 March 2023

How to tackle political polarization — the researchers trying to bridge divides

Political divisions are intensifying, threatening democracies around the world. What strategies bring people closer together again?

- Saima May Sidik



Supporters of former US president Donald Trump climb the west wall of the US Capitol building in January 2021.Credit: Jose Luis

Magana/AP/Shutterstock

On 20 October 2020, in the middle of a bitterly contested US election season, there was an unexpected détente in Utah. Two opposing candidates for governor, Spencer Cox (a Republican) and Chris Peterson (a Democrat), shunned the verbal attacks that dominate most political campaigns. Instead, they released a joint political advertisement in which they made a commitment to remain civil, accept the outcome of the election and eschew the divisive hatred that infests politics today.

“I’m not sure this has ever been done before,” Cox tweeted along with the video ad.

The message went viral. And, according to researchers who later studied its effects, viewing it helped to shore up support for democracy among potential voters (see go.nature.com/3kjgpct).

The video was one of around two dozen interventions tested last year in an initiative called the Strengthening Democracy Challenge, run by social psychologists at Stanford University in California. The team is part of a community of researchers who are trying to find ways to stem the flow of hatred into politics.

Animosity is running high right now. In a survey conducted last year by the Pew Research Center, a think tank in Washington DC, 72% of Republicans said that Democrats are “more immoral” than members of their own party, and 63% of Democrats said the same about Republicans — increases of 17% and 16% in just a three-year period. Similar trends have been seen in other countries[1]. In Switzerland, for example, the degree to which people like their own party more than another has increased by about 60% since the 1980s. The pattern in the United States, however, is especially strong (see ‘Strength of feeling’).

Hostility between political parties has been associated with a lack of respect for democracy and a rise in support for partisan violence[2], such as the attack on the US Capitol in 2021. Some researchers think that these trends could eventually culminate in the collapse of democracy in the United States and elsewhere. “That’s the lurking dark pit that we’re all trying to avoid,” says

social psychologist Kurt Gray at the University of North Carolina at Chapel Hill.

Gray and other social scientists, mostly based in the United States, report that they have made some progress in easing tensions. They have developed ways of engineering conversations so they move away from conflict and towards productive discourse. Some of the most effective methods are time consuming and require a lot of buy-in from participants, so researchers are developing easier interventions — such as watching Cox and Peterson's joint ad — that measurably reduce antidemocratic attitudes, at least in the short term.

At the same time, researchers are beginning to harness technology in ways that might one day bring these interventions to millions of people. And practitioners working at hundreds of bridging organizations — groups that forge links and aid collaboration between people — are turning research into genuine connections between members of opposing groups.

But there are major challenges ahead. Most of these interventions have been used only in small-scale trials, and some researchers doubt that they could be scaled up enough to have a measurable impact. Meanwhile, some studies[3] suggest that politically motivated violence — often thought of as one of the most damaging outcomes of animosity between opposing groups — might be a separate beast altogether, driven by different factors.

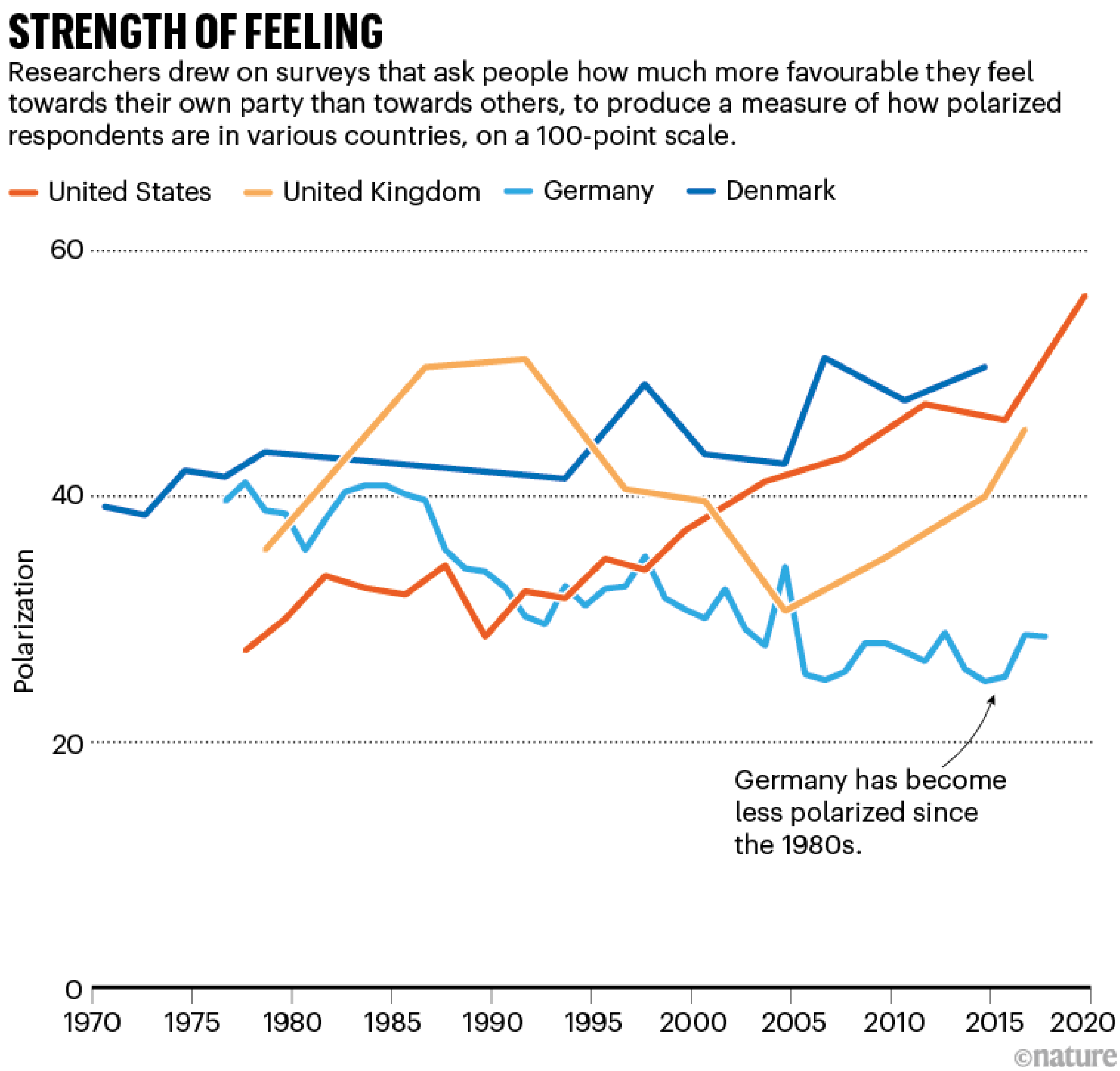
"I think the field as a whole is doing really great work and doing the right kind of work," says political scientist Lisa Argyle at Brigham Young University in Provo, Utah. But changing the tone of an entire nation is not a simple ask. "Right now, we wish we had better answers," she says.

Bitter divide

Hostility between people who support different political parties — or partisan animosity — has increased markedly in the past few decades. In 1978, when people in the United States were asked to rate their warmth towards political parties on a 100-point scale, they rated their own party 27.4 points higher than the other party, on average. By 2020, the gap had widened

to 56.3 points. Several other countries, including Switzerland, France and Denmark, showed similar trends[1], but none to the same degree as the United States — where most research on how to reduce animosity is also taking place.

As trust has diminished, serious threats to democracy have emerged. In the 2022 US midterm elections, more than one-third of Republican nominees running for a state position denied the legitimacy of the 2020 election, according to an analysis by news analysis website FiveThirtyEight (although many of these candidates were not elected).



Source: Ref.1

In the United States, some people seem comfortable with the idea of using violence to hash out political differences. In a study published last year[4], political scientist Sean Westwood at Dartmouth College in Hanover, New Hampshire, and his colleagues found that between 1% and 7% of participants favoured such tactics.

“I’m deeply troubled,” says Westwood. An especially insidious consequence of partisan animosity, in his view, is the way political beliefs are seeping into aspects of life they shouldn’t impact, including hiring practices and even physicians’ advice.

Some researchers blame the proliferation of partisan news sources for the country’s woes. Meanwhile, structural aspects of US politics, such as the two-party system, mean that “us-versus-them is baked into our political process”, says Gray.

These factors breed disagreement, but disagreement alone is not the problem, says Gray. In fact, it is a crucial component of a pluralistic democracy. Problems arise when citizens block out the things that they agree on and begin to focus only on their differences. “Conflict flattens people,” says facilitator John Sarrouf at the bridging organization Essential Partners in Cambridge, Massachusetts.

The study of partisan animosity has its roots in the century-old field of in-group and out-group dynamics. In 1954, US psychologist Gordon Allport outlined a set of conditions that foster acceptance between members of disparate groups, such as having equal status, a common goal and support from authority figures[5]. Many current efforts to reduce partisan animosity still incorporate some or all of these qualities.

In 1988, political scientist James Fishkin, now at Stanford University, stumbled upon one of these methods. At the time, his goal was to find out what voters really thought about the political issues of the day. He devised a system called deliberative polling. Participants from diverse political backgrounds were given briefing materials describing various issues, then left to discuss their views with the help of trained moderators. After the discussion, participants worked together to draft questions for panels of

experts. They alternated between discussions and posing questions to the experts over the course of several days.

It turns out that deliberative polling is a very effective way to bring participants together on polarizing issues[6]. “We get massive changes!” Fishkin says. For example, before a 2021 deliberation exercise involving almost 1,000 people, 35% of Republican participants endorsed the need to reach net zero carbon emissions; after taking part, that figure rose to 55%. Democrats, meanwhile, became more supportive of using nuclear power to reduce greenhouse-gas emissions. Deliberative polling has helped to reduce prejudice against marginalized Roma people in Bulgaria and build trust between Protestant and Catholic people in Northern Ireland[7]. “We have maybe blundered on to a design, but it works,” Fishkin says.

Healing rifts

In the past decade, researchers have deepened their understanding of what contributes to animosity and have learnt lessons about how to apply their knowledge to today’s online culture.

One factor that worsens divides is that people tend to overestimate how different they are from supporters of the opposite political party, imagining their opponents as “simplistic caricatures”, says decision scientist Julia Minson at Harvard University in Cambridge, Massachusetts. Republicans think that around 36% of Democrats are atheist or agnostic (the true figure is about 9%) and Democrats think that around 44% of Republicans earn more than US$250,000 (2% actually do)[8]. In one study[9], helping people with liberal views to correct false beliefs increased their feeling of warmth towards people with a conservative outlook by about 7 points on a 100-point scale.



Supporters of Brazil's former president Jair Bolsonaro break a window of the Supreme Court in Brasilia in January, a week after the inauguration of new president Luiz Inácio Lula da Silva.Credit: On Molina/AFP via Getty

Hoping to foster constructive engagement, Minson and her colleagues used a natural-language-processing algorithm to identify qualities of conversations that make people feel heard[10]. Participants with opposing views held text-based online conversations about contentious issues, then rated how receptive their partners were to their viewpoints. The algorithm correctly predicted which participants would be viewed as receptive by their interlocutors, and which would not.

Study participants who seemed most considerate of others' views hedged their claims, emphasized points of agreement and acknowledged other perspectives, among other behaviours. What's more, Minson has found that she can train people to show these qualities in conversations.

The right framing for a conversation can go a long way. When Sarrouf prepares to facilitate a discussion with a community in conflict, he carefully

considers how to initiate the dialogue. For example, "if I ask you, 'are you for or against gun control?', you've got a polarized conversation right away," he says. Instead, he asks participants to tell a story that can help others to understand how they've reached their beliefs.

Scaling up

Interventions designed for individuals or small groups are inherently difficult to deploy at the scale of a country. Some researchers are designing new interventions — and updating old ones — to have a broader reach.

Fishkin and his collaborators at Stanford's Deliberative Democracy Lab want to bring deliberative polling to the whole United States. The first problem is that "we have to figure out a way where deliberations can run without a human moderator", says Alice Siu, a researcher at the lab. It would be too hard to find enough qualified people to run such a huge number of sessions, even if the money to hire them was available. With the help of Stanford computer scientist Ashish Goel, Fishkin, Siu and their collaborators built a custom online deliberation platform, complete with an automated moderator. In 2021, they used the system in a session based around climate change, and it seems as adept at depolarizing participants as in-person sessions are, says Fishkin.



Inside the US Supreme Court's war on science

Next, they want to scale up. The major hurdle is finding people to take part, Fishkin says. Word of mouth might help, and Fishkin hopes that social-media platforms will help participation to grow. "I think it's going to spread the benefits of deliberation around the world."

But not everyone wants to spend hours discussing politics online, Westwood says. What's more, the positive effects of deliberative polling might wane and need to be reinvigorated once in a while. Fishkin has evidence[6] that the effect on participants' voting habits lasts a year, but would people be willing to take part in multi-day deliberations on an annual basis? "That's a huge ask," says Westwood. Minson suggests testing the system on an intermediate scale before devoting resources to a country-wide trial. "Can we depolarize one city?" she asks.

Social psychologist Robb Willer at Stanford University and his colleagues are studying much less intensive methods for reducing partisan animosity. Their Strengthening Democracy Challenge crowdsourced and tested online interventions — often surveys or videos that already existed — to see whether they reduce anti-democratic attitudes, support for partisan violence or extreme levels of partisan animosity in a group of more than 30,000 online participants.

One of the most successful interventions was an ad for Heineken beer. In the four-minute video, pairs of people with opposing viewpoints work together to achieve a goal (assembling a bar and a set of stools) then sit down to have a drink together while they discuss their views. Watching the ad produced a reduction in partisan animosity of about 10 points out of 100. In a similarly effective intervention, participants read about how the media exaggerate political differences to increase engagement.

Social-media platforms could help to promote these videos, Willer thinks. But even so, getting enough people to watch the videos to make a difference could be tough, Westwood says. Willer suggests a more nuanced approach. What if social-media platforms could use principles identified by research to find naturally occurring content that decreases animosity, such as the Heineken ad, then give it more visibility relative to more polarizing content?

A couple of platforms, he says, have expressed interest in working with him on such projects.

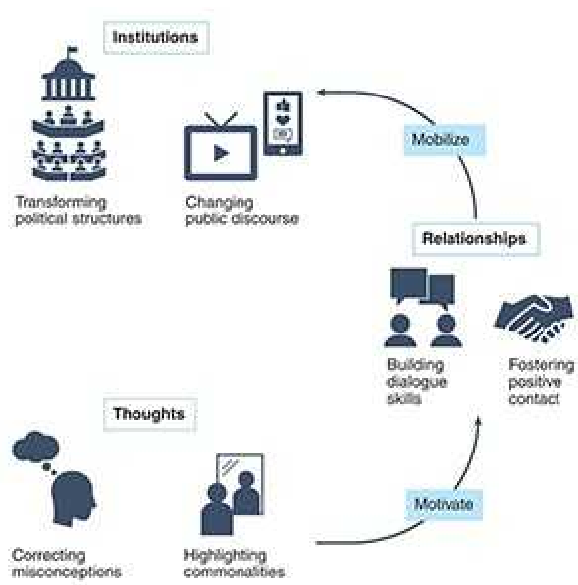
Researchers could randomly select people to view material thought to reduce animosity, and then assess how well it worked.

Argyle is intrigued by this idea, but has questions. For example, “Can we actually identify the kinds of posts that increase partisan animosity and decrease it in a really consistent, reliable way?”

Small solutions to big problems

One major reason for dialling down partisan animosity between individuals is to help counteract bigger problems that threaten democracy. But some research questions whether the two issues are linked. In one study[3], Westwood and his colleagues tested whether a positive experience with a member of the opposite political party could change broader feelings of party loyalty, support for collaboration between parties or support for democratic norms.

They told participants that members of the opposite political party were going to choose to give them some, all or none of $10. In reality, a computer chose whether to give them most or none of the money.



Interventions to reduce partisan animosity

Believing they had a positive experience with the opposite-party member increased the warmth participants felt towards that person, but did little to change their opinion of the wider democratic system, the researchers found.

Willer's lab is currently testing the hypothesis that people who support political violence think in an intrinsically different way from those who don't, and respond to interventions differently.

In terms of research priorities, Minson sees animosity as quite separate from political violence. "I really think we should focus on Grandma and Grandpa at the dinner table and leave the guy with the Viking horns to the prosecutors," she says.

Starting at the top

Some researchers think that visible examples of good behaviour might help to maintain democracy.

Westwood would like to see more politicians promote democratic norms. The joint political ad produced by Cox and Peterson is a prime example of how party elites can model responsible behavioural norms. "And it is as simple as pledging to be positive," he says.

Support from the top has proved crucial in other countries, as well. Peacebuilder Seth Karamage helps to dispel tensions between members of different ethnic and religious groups in Nigeria. When Karamage, who is affiliated with the Karuna Center for Peacebuilding in Greenfield, Massachusetts, starts working in a new region, the first thing he does is to find key politicians whose agendas are well served by reducing violence in the region. Gaining their support increases his access to community leaders.

For the past four years, Karamage has built bridges between people of different religions. A group of farmers, who are Christians, recently told Karamage that last Christmas was the first time they'd celebrated the holiday without enduring an attack from nearby cattle herders, who are mostly Muslims. The impacts of peacebuilding work are "very hard to

quantify”, Karamage says, but anecdotes such as this give him confidence that the approach is working.

“Conflict is the business of human beings,” Karamage says. Even so, he and others are making it their business to reduce the rancour.

Nature **615**, 26-28 (2023)

doi: https://doi.org/10.1038/d41586-023-00573-5

References

1. Boxell, L., Gentzkow, M. & Shapiro, J. M. *Rev. Econ. Stat*. https://doi.org/10.1162/rest_a_01160 (2022).
2. Hartman, R. *et al. Nature Hum. Behav.* **6**, 1194–1205 (2022).
3. Broockman, D. E., Kalla, J. L. & Westwood, S. J. *Am. J. Pol. Sci*. https://doi.org/10.1111/ajps.12719 (2022).
4. Westwood, S. J., Grimmer, J., Tyler, M. & Nall, C. *Proc. Natl Acad. Sci. USA* **119**, e2116870119 (2022).
5. Allport, G. W. *The Nature of Prejudice* (Addison-Wesley, 1954).
6. Fishkin, J., Siu, A., Diamond, L. & Bradburn, N. *Am. Polit. Sci. Rev.* **115**, 1464–1481 (2021).
7. Luskin, R. C., O’Flynn, I., Fishkin, J. S. & Russell, D. *Polit. Stud*. https://doi.org/10.1111/j.1467-9248.2012.01005.x (2012).
8. Ahler, D. J. & Sood, G. *J. Politics* **80**, 964–981 (2018).
9. Reininger, K. M., Krott, N. R., Hoenisch, M., Scheunemann, J. & Moritz, S. *J. Soc. Polit. Psychol.* **8**, 453–472 (2020).
10. Yeomans, M., Minson, J., Collins, H., Chen, F. & Gino, F. *Organ. Behav. Hum. Decis. Process.* **160**, 131–148 (2020).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00573-5

Opinion

- **Will Steffen (1947–2023)** [20 Feb 2023]
  • Earth-system scientist and climate advocate.
- **Autonomous ships are on the horizon: here’s what we need to know** [27 Feb 2023]
  • Ships and ports are ripe for operation without humans — but only if the maritime industry can work through the practical, legal and economic implications first.

- OBITUARY
- 20 February 2023

Will Steffen (1947–2023)

Earth-system scientist and climate advocate.

- Lesley Hughes &
- Martin Rice

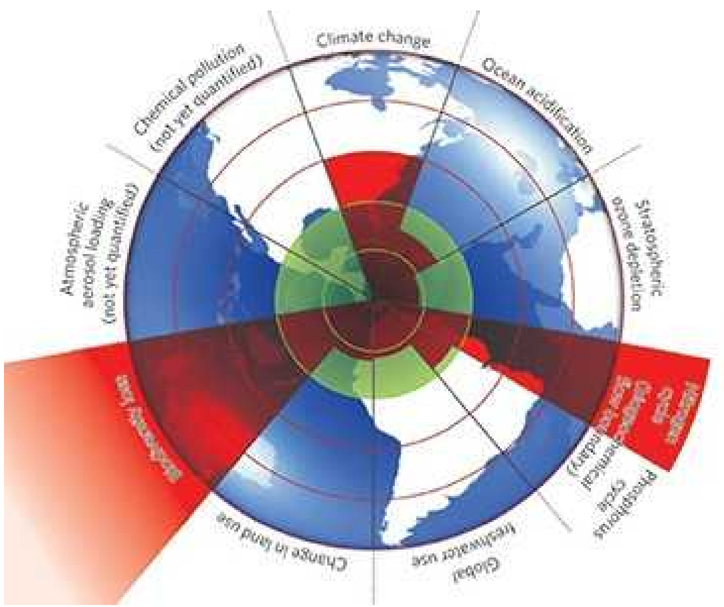


Credit: Alex Hofford/EPA/Shutterstock

Will Steffen was a driving force in Earth-system science as it emerged in the 1980s and 1990s, bringing together disparate fields such as ecology, biology, oceanography and climate science. He was an influential mentor and a passionate public advocate of science-based climate-change policy, speaking truth to power in his adopted home of Australia and worldwide. He has died, aged 75.

Steffen was born in Norfolk, Nebraska, in 1947. Having trained in chemical engineering at what was then the University of Missouri in Rolla and in chemistry at the University of Florida in Gainesville, he emigrated to Australia in the late 1970s to take up a postdoctoral fellowship in X-ray crystallography at the Australian National University in Canberra. He later moved to the Commonwealth Scientific and Industrial Research Organisation (CSIRO), also in Canberra, to work in science management, editing and communication.

Steffen’s passion for science and his exemplary people skills led to him being headhunted for the International Geosphere–Biosphere Programme (IGBP), which aimed at understanding the physical, chemical and biological processes that regulate the whole Earth system. He eventually served as the IGBP’s executive director from 1998 to 2004.



[A safe operating space for humanity]

He was attracted to tackling complex problems in an interdisciplinary way, and with Earth-system science he blossomed as a researcher. Many ideas that he developed and popularized were at first considered radical but are now mainstream. These include the ‘Great Acceleration’[1], which describes the dramatic increase in human environmental impact since the 1950s, brought about by population growth and fossil-fuel burning. He was also a significant advocate of the Anthropocene, the idea popularized by his

colleague, Nobel prizewinner Paul Crutzen, that the planet has entered a new geological epoch because of human activity. Steffen served as the climate-science expert on the Anthropocene Working Group.

Among many seminal publications, Steffen is particularly known for the 2009 *Nature* paper 'A safe operating space for humanity'[2], which introduced the key conceptual framework of 'planetary boundaries'. Along with Johan Rockström and colleagues, Steffen identified nine planet-scale processes that regulate Earth-system stability, together with thresholds that, if crossed, could lead to large-scale environmental tipping points. He also led work on Hothouse Earth[3], exploring how self-reinforcing feedbacks might push Earth into a regime in which its climate could no longer be stabilized even if greenhouse-gas emissions were substantially reduced.

Steffen was fearless in his advocacy of science and science-based policy, communicating the risks of climate change with sobering clarity. In 2011, he was appointed to the Australian government's Climate Commission, which was dedicated to deepening public understanding of climate change and its impacts. When the commission was abolished in 2013, Steffen co-founded the Climate Council of Australia to continue the work, funded by public donations.



Humans versus Earth: the quest to define the Anthropocene

In 2016, Steffen publicly challenged the Australian government after it lobbied the United Nations scientific and cultural agency UNESCO to censor all mentions of impacts on the Great Barrier Reef, the Northern Territory's Kakadu National Park and Tasmania's forests from its report *World Heritage and Tourism in a Changing Climate*. He and colleagues also challenged CSIRO's decision to cut more than 100 climate-science staff. Following international condemnation, the government announced a new CSIRO Climate Science Centre based in Hobart, Tasmania.

Steffen frequently acted as an expert witness in court cases seeking to limit new fossil-fuel developments. In 2022, his evidence helped to block a proposed opencast coal mine in the Galilee Basin in Queensland on the grounds of unacceptable impacts on climate change and human rights, including those of children and First Nations peoples.

Steffen was a kind and generous mentor, and an even-tempered and patient man. But he was scathing in his disdain for "the fossil fuel elite", writing that "these people have no right to destroy my daughter's future and that of her generation". His last major project with the Climate Council, where both of us worked with him, was a report titled *Aim High, Go Fast*. It provided a damning assessment of the inadequacy of the scale and pace of policy action to meet the climate challenge, and reiterated the concept of a "critical next decade" for action.

An outdoors enthusiast and mountaineer, he hiked with his wife, daughter and friends in the Australian wilderness, climbed mountains all over the world and in 2010 published *Himalayan Dreaming*, the definitive history of Australian Himalayan climbing. He commented on the similarities between his passions: "When you come off a big climb, you really appreciate the beauty of what's around you. That's the buzz you get in science when you solve a big problem and suddenly see how it all fits together."

Will Steffen epitomized the ethos of a social contract between scientist and society in his pursuit and sharing of knowledge relevant to the grand challenge of climate change. His visionary academic publications represent a track record of which any scientist would be proud, but his even greater legacy is the thousands of people he educated and inspired to work for a better future.

Nature **615**, 29 (2023)

doi: https://doi.org/10.1038/d41586-023-00519-x

References

1. Steffen, W., Broadgate, W., Deutsch, L., Gaffney, O. & Ludwig, C. *Anthropocene Rev.* **2**, 81–98 (2015).
2. Rockström, J. *et al. Nature* **461**, 472–475 (2009).
3. Steffen, W. *et al. Proc. Natl Acad. Sci. USA* **115**, 8252–8259 (2018).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00519-x

- COMMENT
- 27 February 2023

Autonomous ships are on the horizon: here’s what we need to know

Ships and ports are ripe for operation without humans — but only if the maritime industry can work through the practical, legal and economic implications first.

- Rudy R. Negenborn ,
- Floris Goerlandt,
- Tor A. Johansen,
- Peter Slaets,
- Osiris A. Valdez Banda,
- Thierry Vanelslander &
- …
- Nikolaos P. Ventikos

Show authors



By 2024, the Norwegian container ship Yara Birkeland is expected to carry fertilizer autonomously from plant to port with zero emissions. Credit: Torstein Bøe/NTB/AFP via Getty

In January last year, a Japanese car ferry, the Soleil, became the first large vessel to navigate without human intervention. The 220-metre-long ship automatically berthed and unberthed, turned, reversed and steered itself for 240 kilometres across the Iyonada Sea from Shinmoji in northern Kyushu — manoeuvres that even skilled human operators find challenging.

It is early days, but ships are increasingly deploying sensors and artificial-intelligence (AI) systems to navigate, steer and avoid collisions. As with cars, such advances should improve safety, increase efficiency and — along with cleaner fuels and engines — reduce environmental impacts.

This is crucial: 80% of global trade (around 11 billion tonnes) is transported by sea each year[1]. In 2018, shipping generated around 3% (about 1,000 million tonnes) of global carbon dioxide emissions[2]. The International Maritime Organization (IMO) has committed to halving the industry's greenhouse-gas emissions by 2050.

Seafaring is risky and workers are in short supply. Inefficiencies and congestion at ports add delays and costs, as do accidents, such as the grounding of the container ship *Ever Given* in the Suez canal for six days in March 2021. Streamlining passage through locks, reducing energy consumption and negotiating manoeuvres to avoid collisions would enable safer and more optimal use of waterways.

Some small, fully autonomous boats, typically less than 10 metres long, are already in operation for specialist tasks such as monitoring water quality and infrastructure in the open sea, or as test beds for the technology. But the next couple of years will see a sea change, with the first larger 'maritime autonomous surface ships' planned to start commercial operation.



A cleaner future for flight — aviation needs a radical redesign

Pilot projects include the Norwegian container ship Yara Birkeland, an 80-metre-long vessel that, by 2024, is expected to convey fertilizer autonomously and with zero emissions from a manufacturing plant to an export port. In China, a 120-metre-long electric container ship called Zhi Fei has been demonstrated shuttling under remote (and sometimes autonomous) control between two ports in Shandong province.

In a decade, automated vessels might interact with one another. For instance, the Vessel Train, a pilot project funded by the European Union and coordinated by the Netherlands Maritime Technology Foundation in Rotterdam, uses a crewed lead vessel to head a convoy of smaller,

automated ones that can access small waterways around ports efficiently. Ultimately, fleets of self-steering ships or boats might be managed from maritime traffic-control centres located on shore.

But if autonomous vessels are to fulfil their promise, much remains to be done — and soon. More than 50,000 merchant ships trade internationally, under the flags of some 150 nations. A large, high-tech vessel can cost US$200 million to build, and can operate for decades. Ships are complex technically. They need to work in busy shipping lanes, ports and rough open seas.

Combining maritime systems is daunting — from radar, satellites and GPS, cameras and sensors, to image recognition, data analytics and machine-learning algorithms. And autonomous ships need to be plugged into a broader ecosystem of maritime technologies, including interactions between ships and with cargo handlers, equipment, pilots, traffic services and ports.

Here, we highlight research gaps in six key areas.

Understand the challenges at different levels of autonomy

The roll-out of increasing levels of autonomy needs to be managed to assure safety and to allow regulation to keep up. Operational guidelines are needed for vessels in the four classes of autonomy defined by the IMO.

The first challenge is to add sensors and algorithms based on AI and deep learning to the autopilot systems that are currently used on some crewed ships and boats — classed as IMO autonomy level one. Obstacles such as small boats, debris, swimmers and riverbanks do not appear on radar or on the global ‘automatic identification system’ (AIS) that tracks maritime traffic. Visual and thermal cameras and lasers would give a captain a better view of what is around, helping to avoid collisions, assess risks and plan routes — tasks that are now done manually. But to do that, researchers need to overcome limitations to sensor systems — for instance, that some smaller obstacles can be indistinguishable from waves.



Interactions between humans and autonomous ships will need to be redefined, as for the Chinese ship Zhi Fei, which operates under remote control.Credit: CFOTO/Future Publishing via Getty

Autonomy level two includes remotely controlled vessels that are run by a small on-board crew to act as a back-up and deal with maintenance and cargo. Increasingly used on rivers and canals to transport freight, these are operated at a distance on behalf of ship and boat owners, by maritime technology providers such as Seafar in Antwerp, Belgium, and the US company Sea Machines.

Remotely controlled vessels without crew (level three) are also already in operation. So far, these are mainly ‘drones’ less than 10 metres long, used for measuring water depth or monitoring marine habitats and harbours. Scaling them up to ship size, and carrying cargo, cranes, robotics and fuel, requires passing a higher safety bar, and adding systems to avoid grounding, collisions and loss of communication.

Some small drones have achieved full autonomy (level four). In these, the operating system makes decisions and determines actions by itself for a

period. For example, the Wave Glider, AutoNaut, Sailbuoy and Saildrone boats can operate independently for days in the open ocean, powered by wind, solar and wave energy. The challenges of operating near coasts, among other vessels and in shallow water and currreents remain to be addressed.

For vessels at all levels of autonomy, guidelines must also encompass mixed environments in which autonomous and crewed vessels share the same waters. Researchers need to develop cooperative navigation and communication systems in and between groups of vessels, enabling them to operate as a fleet.

Define the role of humans

Autonomous ships will always have a human somewhere in the loop, to check on navigation, perform maintenance, handle cargo, supervise and monitor tasks, and gauge risks. These roles must be carefully specified. Until any ship with a high level of autonomy has been fully tested, at least one human should retain command and control. They might be on board, like an airline pilot, or remote, like a drone operator. As systems improve, human supervision might be required only in emergencies.

Researchers need to design human–machine interfaces to support decision-making for navigation, remote control and interactions with people on other ships and at ports. Remote-control centres will look and feel like those used to direct air traffic. A few centres are already in place, including those running the Zhi Fei, Yara Birkeland and Seafar vessels.



Uncrewed Saildrone Explorer boats can operate autonomously for days in the open ocean, powered by wind and solar energy.Credit: Mass Communication Specialist 2nd Class Dawson Roth/AP/Shutterstock

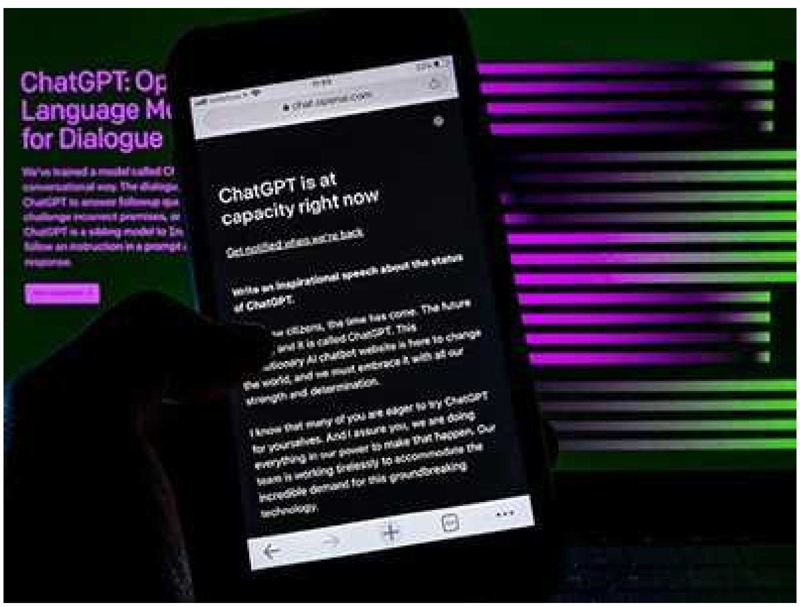
More analysis is needed of how humans and AI interact, including how best to relay performance or navigation information to and from remote operators. Methods need to be designed for verifying automated detections by sensors of small boats and other hazards. And more needs to be learnt about how people understand and anticipate the manoeuvres of other ships to avoid collisions in busy waters. Humans might need to check that the calculated risks of voyages are acceptable in light of weather forecasts and other uncertainties.

Assure safety and security

As in the automotive sector, development of maritime technologies must have safety at the centre. Preliminary guidelines for using automated processes for navigation and systems maintenance have been published by classification societies such as DNV in Norway and the French firm Bureau

Veritas (see go.nature.com/43kxqte and go.nature.com/43me6e7). These guidelines cover processes for qualification of concepts and technologies, and how systems supporting the autonomous and remote operation of vessels should be designed. But they lack specifics about how they should be applied, for example, in poor visibility, during storms or in sea ice.

Ships have a lot of moving parts (engines, generators, propellers, cranes and hatches) that require observation and maintenance by humans for safe operation. Researchers need to develop smart maintenance procedures, which monitor components and identify, diagnose and repair faults remotely. More redundancy in systems, with spare components available to take over when one has failed, would increase resilience.



ChatGPT: five priorities for research

Current guidelines also say little about cybersecurity risks, which will increase in the context of autonomous ships. In the past few years, cyberattacks on major shipping companies — such as the Danish firm Maersk in 2017 and on South Korea's HMM and Japan's K Line in 2021 — have damaged assets and finances. Maersk was forced to rebuild its IT infrastructure in 10 days and sustained losses of more than $300 million, as well as reputational damage. Fleets came to a standstill, blocking ports and delaying cargoes.

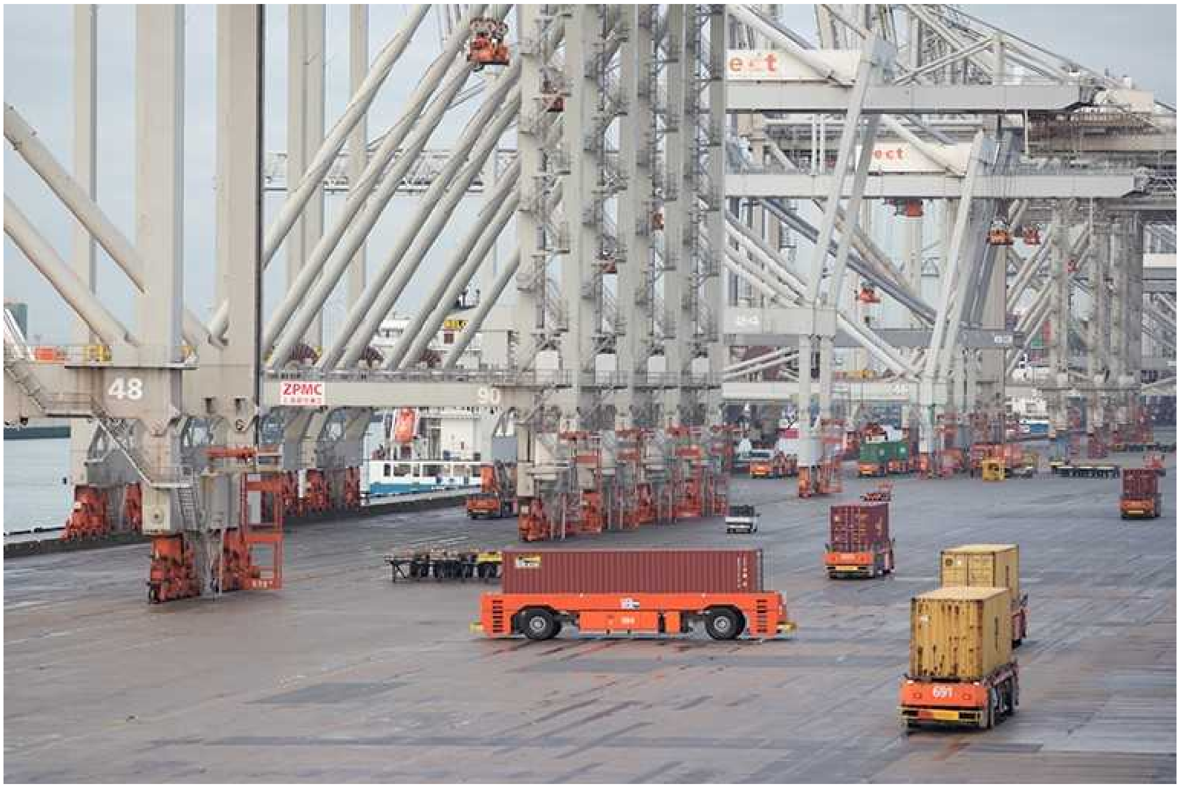
To extend the guidelines, researchers should define safety and security requirements for autonomous ship technologies in a range of operational contexts. For example, how should human–machine interfaces be assured to work when a ship is rolling on heavy seas? How should uncertainties in navigation be handled when operating in currents, winds and tides? Sensors need to distinguish different types of sea ice and feed that back to ship systems to find the safest route. Measures for preventing and responding to cyberattacks need to be developed.

Infrastructure such as buoys, antennas and IT systems at sea and ashore need to be secure, and data links between ships and control centres reliable. The bandwidth of wireless communication systems is influenced by environmental and weather conditions. Digital twins (computer-based copies of large systems) are helpful for monitoring, verifying and validating functional and safety requirements for autonomous ships through simulation. Current models accurately replicate ship navigation by combining data on water depths, sea-bed composition, tidal heights and water visibility, as well as land and weather data.

Rethink ports

Autonomy does not demand radical changes to ship design, although advances will be needed to accommodate equipment to support autonomous operation and the parallel development of cleaner propulsion systems.

It is a different matter for ports, where the advent of autonomous ships will accelerate trends towards fuller automation. The current focus is on automating cargo handling — in 2020, more than 800 million containers were moved around the world by human-operated cranes and vehicles. For example, in the port of Rotterdam, unmanned cranes and ‘automated guided vehicles’ allow an entire container terminal to be operated by 10–15 people each day. Robotic mooring and crane systems are in use in Stockholm, Tallinn and in the Finnish ports of Naantali and Helsinki. Singapore is constructing the world’s largest autonomous terminal at a cost of around US$15 billion, which is expected to be completed in 2040.



Automated guided vehicles move shipping containers in the port of Rotterdam, the Netherlands.Credit: Jasper Juinen/Bloomberg via Getty Images

Autonomous ships will require more services, including automated pilotage and tug assistance, arrival management and berth allocation. Pilots, who now board ships to help crew to enter ports, and 'vessel traffic services', which monitor traffic and provide recommendations to avoid hazards, will need to take control of the vessel remotely. Ways to connect ropes between tugs and crewless ships will be needed. Further ahead, autonomous ports will need to interact with one another and along logistics chains, to ensure that cargo flows seamlessly without bottlenecks and delays. Integrating all the different systems is a major challenge.

Embed autonomy in legal and regulatory frameworks

Maritime trade is heavily regulated through a web of international rules negotiated over centuries by many countries with many legal systems. The

United Nations Convention on the Law of the Sea (UNCLOS) sets out the rights and responsibilities of states for vessels under their jurisdiction. The framework seems well equipped to incorporate autonomous ships. But there are challenges.

Long-established legal concepts underlying maritime law must be reconsidered. For instance, the concepts of 'seaworthiness' (the condition of the vessel and the ability of crew to operate it), 'master' (a single person in charge of a vessel) and the 'genuine link' (between the vessel and the state conferring its nationality on that vessel through its flag) no longer hold for autonomous vessels. Whether the nationality of a vessel is a suitable basis for regulating autonomous ships needs to be questioned, given that their operation can involve many actors spread around the globe.



Cooperative AI: machines must learn to find common ground

The IMO has begun to develop a code for autonomous ships[3], although it has yet to set goals, requirements and rules. Initially, the code will be non-mandatory, potentially becoming mandatory in 2028. It will aim to balance the benefits of these technologies against safety and security concerns, the impact on the environment and on international trade, the potential costs to the industry and the impact on personnel on board and ashore.

Key research questions include: how to design the land-based communications and smart navigation systems that ships interact with. How can the code be harmonized with other national and international regulations — for instance, those related to equipment in ports and land-based logistics? How can conformity with the code be verified? What skills and training are needed for operators in remote-control centres?

Legal researchers should address civil liability, for when third parties incur damages caused by autonomous vessels. Currently, those affected need to prove a wrongful action or omission by a person or by the ship. It is hard to hold AI accountable. Interpretations will also vary across jurisdictions. If companies cannot assess their liability exposure, they might not invest in autonomous vessels and infrastructures. The European Commission's development of an AI Liability Directive offers a positive example, because it will lay down rules for operating markets and standards of proof for claims involving AI-based systems.

Set out the case for autonomous ships

The economic costs and benefits of ship automation and autonomy need to be established if the technology is to be taken up. Balancing all the factors across many parties is challenging, however. It is hard to know how many crew members will be required, on board or on shore. The impacts of illness, strikes and technical problems, and thus rates of shutdowns, must be judged.

More evidence is also needed regarding the economic, environmental and safety implications of autonomous vessels. Equity is key: who will reap the benefits and who will bear the costs? The societal costs of lost crew jobs must be counted, but automated shipping will also require higher-skilled staff. Redundancy payments and costs of hiring or training staff need to be factored in, from business and welfare perspectives. The existing workforce will need reskilling, with new training opportunities put in place — across naval architects, maritime engineers, ship crew and remote-control operators, port and shipping managers and maritime administrators.

Optimizing operations and logistics chains might deliver savings, yet managing vast quantities of data adds expense. Investments will be needed

for archiving and sharing data securely, and for building and staffing data centres.

Inequitable distribution of benefits and costs of innovation can lower public trust in regulators and politicians. The concentration of economic power in the large corporations that might control global autonomous logistics chains must be looked at: for example, food security could be affected if a few multinationals control key supply chains.

Unlike for self-driving cars, there is scant research on how the public perceives risk around autonomous vessels. More studies are needed, especially in contexts in which world views might conflict. For instance, in Canadian coastal areas, Indigenous peoples have a right to be involved in decision-making on matters related to protection of the marine environment.

Autonomous ships are already on the horizon. For them to really make a difference, scholars need to improve understanding of how waterborne autonomy can be adopted.

Nature **615**, 30-33 (2023)

doi: https://doi.org/10.1038/d41586-023-00557-5

References

1. United Nations Conference on Trade and Development. *Review of Maritime Transport* (UNCTAD, 2022).

2. International Maritime Organization. *Fourth IMO Greenhouse Gas Study* (IMO, 2020).

3. ABS. *News Brief: MSC 106* (ABS, 2022).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00557-5

Work

- **Hyperauthorship: the publishing challenges for ‘big team’ science** [27 Feb 2023]
  • Studies involving hundreds, even thousands, of scientists are on the rise, but how do such large groups coordinate their work?
- **I brokered a historic deal to protect the seas** [27 Feb 2023]
  • Julie Robinson, a marine ecologist-turned-policymaker in Belize, is helping to save her country through seaweed farming.

- CAREER FEATURE
- 27 February 2023

Hyperauthorship: the publishing challenges for 'big team' science

Studies involving hundreds, even thousands, of scientists are on the rise, but how do such large groups coordinate their work?

- Bianca Nogrady

Find a new job



Research at multinational facilities such as CERN, Europe's particle-physics laboratory, can lead to studies with huge author lists.Credit: Adam Hart-Davis/SPL

The existence of the Higgs boson was first posited in a trio of papers in 1964. Two of those[1,2] were authored solely by UK theoretical physicist Peter Higgs and the other[3] was co-authored by his US and Belgian counterparts Robert Brout and François Englert.

Nearly half a century later, the experimental confirmation that the Higgs field existed was published in a paper[4] with 2,932 authors. Three years after that, a paper[5] detailing a more accurate measurement of the mass of the Higgs boson set a new record for the highest number of authors on a single paper: 5,154.

Then the COVID-19 pandemic broke that record, with 15,025 co-authors on a research paper[6] examining the effect of SARS-CoV-2 vaccination on post-surgical COVID-19 infections and mortality.

The term 'hyperauthorship' is credited to information scientist Blaise Cronin[7] at Indiana University in Bloomington, who used it in a 2001 publication to describe papers with 100 or more authors. But with the rise of large international and multi-institutional scientific collaborations — such as the ATLAS consortium behind the discovery of the Higgs boson — papers with hundreds, even thousands, of authors are becoming more common. There are many legitimate reasons for this shift, but it is raising questions — and concerns — about the nature of authorship and the impact that hyperauthorship has on the metrics of scientific achievement.

A 2019 report from the Institute for Scientific Information (ISI), part of citation-analytics company Clarivate, noted that the number of papers with more than 100 authors doubled between 1998 and 2011, from around 300 to around 600 (see go.nature.com/3I5j8zt). When the authors of the report looked at how the rates of different levels of hyperauthorship had changed between two five-year time periods — 2009 to 2013 and 2014 to 2018 — they saw that the greatest increase was in articles with more than 1,000 authors.

It’s a big change from the 1980s and is driven largely by an increase in international collaboration in science, says Jonathan Adams, chief scientist at the ISI and co-author of the report. In the 1980s, most scientific papers — more than 90% — listed authors from a single country. That changed around the mid-2000s, with a notable shift to bilateral collaborations between researchers from two nations, Adams says, and “as soon as you get bilateralism, of course the authorship count inevitably goes up”. Then came a rise in collaborations involving multiple nations, with a particular increase in papers involving up to 30 countries. The expansion of multi-authorship and hyperauthorship isn’t limited to particular countries or driven by national policies. “It’s simply the way in which science and research generally has changed,” he says.



Unearned authorship pervades science

An analysis of data in the Nature Index, which tracks research publications in 82 high-quality natural-science journals, also reveals a substantial increase in certain fields in the number of papers with 50 or more authors. For example, in medical and health sciences, the number of such papers rose from 58 in 2015 to 203 in 2021, an increase of 250%; this included a 90% increase from 2019 to 2020, the first year of the COVID-19 pandemic. Physical sciences, meanwhile, was the most common field of research for these multi-authored papers, with 335 articles in 2021, an 18% increase on 2015.

These changes reflect the growing need for large research groups, spread across different types of institution and geographies, to answer complex questions. They also reflect a desire for more-inclusive authorship that recognizes researchers from backgrounds that might have been overlooked in the past. The changes might also be the result of funding from bodies such as the European Research Council, which aids and encourages multinational collaborations.

But hyperauthorship creates challenges for researchers and for the journals that publish their work. Coordinating so many individual contributions across a multitude of institutions and nations is an enormous logistical feat. And hyperauthorship is raising philosophical questions about what it means to be an author of a research paper, and who has the right to — and need for — acknowledgement.

The need for statistical power

A key problem in studies that explore the role of genetics in mental-health disorders is statistical power, says Sarah Medland, a psychiatric geneticist at the QIMR Berghofer Medical Research Institute in Brisbane, Australia, and that's one reason that her field lends itself to hyperauthorship. To examine the effects of genetic variants on the structure and function of the brain, she and her colleagues often use genome-wide association studies — in which large numbers of genomes are scanned for genes that correlate to particular diseases or traits — as well as magnetic resonance imaging (MRI) and psychiatric assessments. But the neurological effect of a single genetic variant can be extremely small, so studies need a lot of participants to find whether effects are statistically significant and replicable. The bigger the studies, the more expensive they are, especially when MRI is involved.

Frustrated by this lack of statistical power, Medland and her colleagues formed ENIGMA — Enhancing Neuro Imaging Genetics through Meta Analysis — a consortium that brings together researchers in imaging genomics, neurology and psychiatry from 50 working groups around the world.

ENIGMA generates papers that are meta-analyses of unpublished data from individual research teams within those working groups. “All of those groups could write their own paper, but if they did that they would be polluting the literature with all of these underpowered studies,” Medland says. Instead, they contribute those unpublished data, or a summary of them, to the meta-analyses. The end result is studies that have hundreds of contributors but are large enough to generate statistically significant and meaningful findings.



[‘Disruptive’ science has declined — and no one knows why]

Hyperauthorship such as this comes about through good communication, both online and in person, says Adams. It often starts at international conferences, where researchers realize that their closest potential collaborators “are certainly not people down the corridor” of their institution, and “probably not people in the same country”, but instead are scientists in their field working abroad. “At the same time, because we’ve got all of these new approaches to communication, we have the ability to bring together very large data sets.”

A desire for more inclusivity in science is also contributing to longer author lists, Medland suggests. “What I think happens if you say you can only have four people” listed on a paper, she says, is that the people who remain as authors tend to be the most senior. “I would suspect that our author lists are

younger, more female and more diverse," now that more authors can be included, she adds.

Hyperauthorship is also the result of scientists in some fields seeking answers that require not just large-scale collaboration, but also huge resources and equipment. This is particularly evident in high-energy physics, where the cost of equipment such as particle accelerators can run into billions of dollars.

No single nation can afford that price tag, says Michael Thoennessen, a nuclear physicist and editor-in-chief at the American Physical Society in College Park, Maryland. "If you want to stay at the forefront and do the most exciting physics, and do things that nobody's ever done before, it gets more and more expensive," he says. That means multinational collaborations, and potentially hundreds or thousands of authors.



Heidi Baumgartner says quantifying authorship is the philosophical elephant in the room.Credit: Jeanina Casusi

The increasing length of author lists, however, does lead to the question of what level of contribution entitles researchers to be included. Heidi Baumgartner, executive director of the ManyBabies consortium — an international, multilaboratory collaboration in developmental psychology research — says that the project uses what's known as the CRediT, or Contributor Roles Taxonomy, to determine what qualifies someone for authorship. This taxonomy was developed by the National Information Standards Organization in the United States; it describes 14 roles that are commonly found in a research context, ranging from conceptualization and funding acquisition to supervision and writing.

"We say you have to make a contribution to at least one of those categories, plus reviewing the manuscript, and that is what merits authorship," says Baumgartner, who is also a social-science research scholar at Stanford University, California. This expectation is clearly laid out at the start of any project in a collaboration agreement, and it ensures that everyone who gives their time and effort is credited equally.



'Ignored and not appreciated': Women's research contributions often go unrecognized

It's not always so straightforward, however. Nicholas Coles, an experimental psychologist at Stanford University and director of the Psychological Science Accelerator, a global network of psychological-science labs, says

that although scientific contributions have been under-recognized in authorship, there is also a well-documented history of an abuse of authorship practices resulting in over-recognition. He refers to “helicopter authorship”, where researchers “recruit large teams to collect data, and then never once acknowledge their contribution”.

This cuts to a bigger question of what sorts of contribution to scientific research warrant authorship — which Coles calls the “coin” of the scientific realm — rather than just a mention in the acknowledgements section at the end of the paper.

Baumgartner says this is the “philosophical elephant in the room of academia”: “How do you actually quantify what authorship means in terms of what contribution someone made, and then how that is used as a currency in terms of what it means for hiring or promotions?” Should the technician who keeps the lab equipment and computers running, or the nurse who collects blood samples, be listed as an author?

Thoennessen says that, in his experience, authorship is limited to scientific staff, who are the ones who are evaluated and rewarded on the basis of their scientific output.

Complex coordination

Getting a research paper published typically requires a large amount of work, even for papers with just a few authors, let alone several thousand. To reduce the risk of such a large ship foundering, ManyBabies uses a form of early peer review called Registered Reports to evaluate the viability of a research project before the data has even been collected. The initiative is being promoted by the Center for Open Science, a non-profit technology organization based in Charlottesville, Virginia, to encourage best practice in scientific methods. Journals, such as *Nature*, that have adopted the initiative commit to publishing a paper if the question and methodology meet the required standard (see *Nature* **614**, 594; 2023).

Baumgartner says that using Registered Reports “gives our contributors, who are all over the world, a little assurance that sticking with this project is

going to result in something, as opposed to the typical process where you run a study and you don't really know if it's ever going to find a home somewhere".

Further challenges in coordinating large authorship groups arise once a study's results are collated and analysed, and the process of writing and editing the manuscript begins. "What happens is the core team that's worked on the meta-analysis will usually prepare the draft and that usually involves both the senior and the first and the last authors," Medland says. "Once there's a draft they're happy with, it gets opened for comments."

Both ENIGMA and ManyBabies use Google Docs to share the draft paper with all co-authors and collate their comments and edits. "Google Docs is both good and bad in that way", says Medland, because you can end up with multiple, lengthy comment threads. Baumgartner says that the Google Docs file can get a little chaotic, "but it's kind of fun to be in a Google Doc where a bunch of other people are also editing it and you can see where someone's working and follow their thought process and then jump in after somebody else".

The next challenge is ensuring that all the authors' affiliations and conflicts of interest are correct. Coles says that one useful tool in this respect is a web app called tenzing, developed by an international team of researchers, which essentially takes a spreadsheet with author names, affiliations and contributions and "writes your authorship page".

Dealing with so many authors is also a headache for publishers. American Physical Society journals put the onus on authors to verify that they are contributors to a paper. Its system "automatically sends an e-mail to those authors saying, 'the paper has been submitted with your name on it — unless we hear from you, we take this as an agreement that yes, you contributed and you are co-author,'" Thoennessen says.



[African researchers lead campaign for equity in global collaborations]

How authorship is eventually structured in the journal varies enormously between fields and even subfields. “We actually have what’s called a first authorship group and then we have an analyst group, and so on. And within the groups, usually it’s alphabetical order by surname,” Medland says. A multi-authorship paper covering a genome-wide association study might therefore have several lists of authors, each ordered alphabetically. By contrast, the hyperauthored Higgs boson paper has a single alphabetically ordered list.

Differences in authorship conventions can create problems when it comes to researchers getting appropriate recognition for their contributions over their career. “As an early-career researcher working in this type of field, where you have large authorship, you often run into statements like, ‘You didn’t do anything, because you’re a middle author,’” Medland says. “Authorship conventions in general are really misunderstood.” She contrasts medical research, in which the last author is considered the most senior, with psychology, where the person in that position has made the smallest contribution.

Thoennessen would like to see standards for authorship that apply across all fields, to ensure that individual researchers aren’t disadvantaged by their discipline’s conventions.

Overweight impact

When 15,000 or so researchers have contributed to a paper, it distorts the metrics commonly used to evaluate the impact and importance of a particular research project.

These papers often do represent “something that’s closer to the cutting edge of science than the single-author paper”, says Adams. But their “outrageous” citation count is also the result of the ‘two home crowds’ effect, whereby all the nations represented in one of these hyperauthored papers — and all the authors — are going to trumpet their achievements and thereby drive a higher rate of citations. This can also have the effect of distorting the metrics of scientific achievement for the nations and institutions involved, particularly if they are smaller fish in the global scientific pond.

Adams suggests that the number of authors on publications should be considered when looking at citation counts for institutions or nations, just as citation counts are normalized for year of publication and field. “We should be normalizing for the authorship, so you get a more representative citation indicator.” He says there’s even a case for leaving the largest of hyperauthored papers out of the citation process entirely. “If it’s CERN or one of the big telescopes, it’s like you’re either in the club or you’re not,” Adams says. “And then, what are we comparing you with?”

Despite the increase in the number and profile of hyperauthored papers, there’s still a lot of uncertainty about how to deal with them. Coles says that some journals are cautious about hyperauthored papers, whereas others welcome them. But as ‘big team’ science becomes more common and more respected, it might force a reckoning with what these papers mean for the scientific enterprise.

“The more authors that you’re working with, the more complicated things get,” Coles says. “That requires some pretty new thinking, from both researchers and journals, and the people who evaluate science.”

Nature **615**, 175-177 (2023)

doi: https://doi.org/10.1038/d41586-023-00575-3

References

1. Higgs, P. W. *Phys. Lett.* **12**, 132–133 (1964).
2. Higgs, P. W. *Phys. Rev. Lett.* **13**, 508–509 (1964).
3. Englert, F. & Brout, R. *Phys. Rev. Lett.* **13**, 321–323 (1964).
4. Aad, G. *et al. Phys. Lett. B* **716**, 1–29 (2012).
5. Aad, G. *et al. Phys. Rev. Lett.* **114**, 191803 (2015).
6. COVIDSurg Collaborative & GlobalSurg Collaborative *Br. J. Surg.* **108**, 1056–1063 (2021).
7. Cronin, B. *J. Assoc. Inf. Sci. Technol.* **52**, 558-569 (2001).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00575-3

- WHERE I WORK
- 27 February 2023

I brokered a historic deal to protect the seas

Julie Robinson, a marine ecologist-turned-policymaker in Belize, is helping to save her country through seaweed farming.

- Chris Woolston



Julie Robinson is the director of the Nature Conservancy’s Belize Program in Punta Gorda, Belize. Credit: Meridith Kohut

In Belize, seaweed is a lifeline. In this picture, I'm holding a chunk of *Eucheuma isiforme*, a seaweed harvested from an underwater farm in the Turneffe Atoll, about 40 kilometres off the coast. As a little girl growing up in Belize City, I remember adults drinking 'seaweed shakes', creamy drinks often served with rum. The shake is tasty and loaded with minerals, and it has a reputation as an aphrodisiac. The seaweed was so sought after that it practically disappeared from the wild. Now, seaweed farms provide locals with a sustainable source of income. Seaweed is easy to grow, and the biomass can multiply tenfold in the space of months.

The clumps of seaweed are like an artificial reef that brings new life to the waters around the atoll. The branches provide habitats for juvenile lobsters, the single most economically important fishery species in the Caribbean. Small invertebrates attract parrotfish and snappers, and those fish bring barracuda and other predators. Tour operators will visit the farms just for the wildlife.



Building a living shoreline to help combat climate change

As director of the Nature Conservancy's Belize Program, I helped to negotiate a deal with the Belize government in 2021 that allowed the country to exchange US$364 million of debt for a commitment to protect 30% of its ocean waters. About half of those protected waters will be completely closed to fishing. Now, some fishers can earn a living from seaweed farming. It was

the largest debt swap for marine conservation in history, and it happened in little Belize. It was transformative for the country. I’ve always wanted to do something positive for my home, and I know that I’ve accomplished that goal.

I’m trained as a marine ecologist, but now that I work on the policy side, I spend most of my time in front of my computer. Still, I will go out to meet with partners or donors to showcase the seaweed farms and our other projects. That’s how I get energized. I’m happiest when I’m in the water.

Nature **615**, 180 (2023)

doi: https://doi.org/10.1038/d41586-023-00576-2

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00576-2

| Section menu | Main menu |

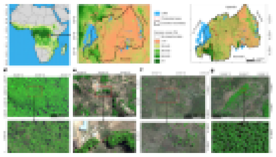
Research

- **Sub-continental-scale carbon stocks of individual trees in African drylands** [01 Mar 2023]
  • A database and viewer is described, resulting from the assessment of the carbon stock of over 9 billion individual trees in semi-arid sub-Saharan Africa using field data, machine learning, satellite data and high-performance computing.
- **Oceanic climate changes threaten the sustainability of Asia's water tower** [01 Mar 2023]
  • Weakening blocking effect of the High Mountain Asia on the westerlies-carried deficit in precipitation minus evaporation from the southeast North Atlantic is demonstrated, leading to persistent northward expansion of terrestrial water storage deficit in the Tibet Plateau.
- **Annelid functional genomics reveal the origins of bilaterian life cycles** [25 Jan 2023]
  • Comparative chromosome-scale genome sequencing and transcriptomic and epigenomic profiling of three different species of annelid provide insight into the evolutionary origin of larvae.
- **Essential elements of radical pair magnetosensitivity in Drosophila** [22 Feb 2023]
  • Non-CRYPTOCHROME-dependent radical pairs can elicit responses to magnetic fields in neurons in Drosophila melanogaster.
- **Palaeogenomics of Upper Palaeolithic to Neolithic European hunter-gatherers** [01 Mar 2023]
  • Combined analysis of new genomic data from 116 ancient hunter-gatherer individuals together with previously published data provides insights into the genetic structure and demographic shifts of west Eurasian forager populations over a period of 30,000 years.
- **FXR inhibition may protect from SARS-CoV-2 infection by reducing ACE2** [05 Dec 2022]
  • FXR regulates the levels of ACE2 in tissues of the respiratory and gastrointestinal systems that are affected by COVID-19, and inhibiting FXR with ursodeoxycholic acid downregulates ACE2 and reduces susceptibility to SARS-CoV-2 infection.
- **Spike and nsp6 are key determinants of SARS-CoV-2 Omicron BA.1 attenuation** [11 Jan 2023]

• Comparisons of a chimeric recombinant version of SARS-CoV-2, the ancestral SARS-CoV-2 strain and an Omicron isolate show that the attenuated virulence of Omicron is caused by mutations in both the spike protein and nsp6.

- ## Active eosinophils regulate host defence and immune responses in colitis [12 Dec 2022]
  • Single-cell transcriptomic profiling and functional assays are used to identify subpopulations of eosinophils that are present in the mouse gastrointestinal tract and provide insight into the role of these cells in inflammatory bowel diseases in humans.

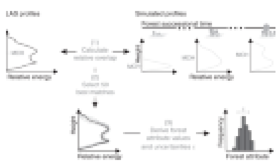
- ## Microbiota-derived 3-IAA influences chemotherapy efficacy in pancreatic cancer [22 Feb 2023]
  • Indole-3-acetic acid (3-IAA), a tryptophan metabolite derived from the gut microbiota, is associated with a better response to chemotherapy in pancreatic ductal adenocarcinoma (PDAC), and dietary interventions could have a role in the treatment of PDAC.



Nation-wide mapping of tree-level aboveground carbon stocks in Rwanda

22 December 2022

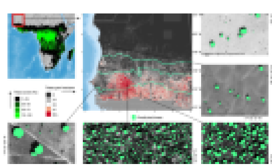
Maurice Mugabowindekwe, Martin Brandt, … Rasmus Fensholt



From small-scale forest structure to Amazon-wide carbon estimates

08 November 2019

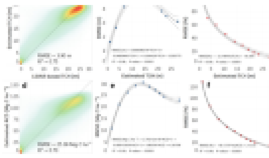
Edna Rödig, Nikolai Knapp, … Andreas Huth



An unexpectedly large count of trees in the West African Sahara and Sahel

14 October 2020

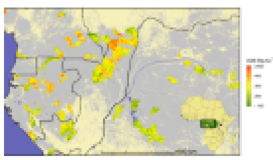
Martin Brandt, Compton J. Tucker, … Rasmus Fensholt



[Monitoring tropical forest carbon stocks and emissions using Planet satellite data]

28 November 2019

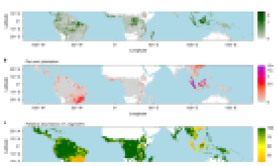
Ovidiu Csillik, Pramukta Kumar, … Gregory P. Asner



[A map of African humid tropical forest aboveground biomass derived from management inventories]

08 July 2020

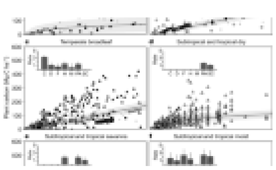
Pierre Ploton, Frédéric Mortier, … Sylvie Gourlet-Fleury



[The expansion of tree plantations across tropical biomes]

06 June 2022

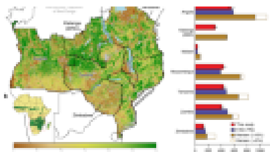
Matthew E. Fagan, Do-Hyung Kim, … Elsa M. Ordway



[Mapping carbon accumulation potential from global natural forest regrowth]

23 September 2020

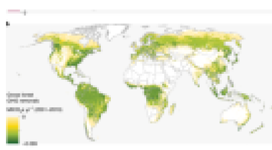
Susan C. Cook-Patton, Sara M. Leavitt, … Bronson W. Griscom



[Carbon losses from deforestation and widespread degradation offset by extensive growth in African woodlands]

02 August 2018

Iain M. McNicol, Casey M. Ryan & Edward T. A. Mitchard



[Global maps of twenty-first century forest carbon fluxes]

21 January 2021

Nancy L. Harris, David A. Gibbs, … Alexandra Tyukavina

- Article
- Open Access
- [Published: 01 March 2023]

Sub-continental-scale carbon stocks of individual trees in African

drylands

- Compton Tucker ,
- Martin Brandt ,
- Pierre Hiernaux ,
- Ankit Kariryaa,
- Kjeld Rasmussen,
- Jennifer Small,
- Christian Igel,
- Florian Reiner,
- Katherine Melocik,
- Jesse Meyer,
- Scott Sinno,
- Eric Romero,
- Erin Glennie,
- Yasmin Fitts,
- August Morin,
- Jorge Pinzon,
- Devin McClain,
- Paul Morin,
- Claire Porter,
- Shane Loeffler,
- Laurent Kergoat,
- Bil-Assanou Issoufou,
- Patrice Savadogo,
- Jean-Pierre Wigneron,
- Benjamin Poulter,
- Philippe Ciais,
- Robert Kaufmann,
- Ranga Myneni,
- Sassan Saatchi &
- …
- Rasmus Fensholt

Show authors

Nature **volume 615**, pages 80–86 (2023)

- 3315 Accesses
- 213 Altmetric
- Metrics details

Abstract

The distribution of dryland trees and their density, cover, size, mass and carbon content are not well known at sub-continental to continental scales[1,2,3,4,5,6,7,8,9,10,11,12,13,14]. This information is important for ecological protection, carbon accounting, climate mitigation and restoration efforts of dryland ecosystems[15,16,17,18]. We assessed more than 9.9 billion trees derived from more than 300,000 satellite images, covering semi-arid sub-Saharan Africa north of the Equator. We attributed wood, foliage and root carbon to every tree in the 0–1,000 mm year^{-1} rainfall zone by coupling field data[19], machine learning[20,21,22], satellite data and high-performance computing. Average carbon stocks of individual trees ranged from 0.54 Mg C ha^{-1} and 63 kg C tree^{-1} in the arid zone to 3.7 Mg C ha^{-1} and 98 kg tree^{-1} in the sub-humid zone. Overall, we estimated the total carbon for our study area to be 0.84 (±19.8%) Pg C. Comparisons with 14 previous TRENDY numerical simulation studies[23] for our area found that the density and carbon stocks of scattered trees have been underestimated by three models and overestimated by 11 models, respectively. This benchmarking can help understand the carbon cycle and address concerns about land degradation[24,25,26,27,28,29]. We make available a linked database of wood mass, foliage mass, root mass and carbon stock of each tree for scientists, policymakers, dryland-restoration practitioners and farmers, who can use it to estimate farmland tree carbon stocks from tablets or laptops.

Main

Improved knowledge of dryland trees, defined here as having a green crown area >3 m^2 with an associated shadow (Extended Data Fig. 1), is essential to

understand their roles in local livelihoods, economies, ecosystems, the global carbon cycle and the climate system in general. Basic information about the distribution of dryland trees and their density, cover, size, mass and carbon content are not well known[2,3,4,5]. This knowledge is required for understanding the functional traits of trees in relation to water resources with changes in climate, predicted increase in aridity and the number and duration of drought events[30,31]. The sources of information used to estimate carbon stocks in drylands include field surveys at plot scale; ecosystem models[23]; and low-resolution, moderate-resolution and high-resolution satellite images[4,5,6,7,8,9,10,11,12,13,14], which are used to infer bulk properties such as averages of tree cover, dry masses and carbon density per unit area at a much coarser spatial scale than individual trees.

Although most emphasis is put on the development of advanced monitoring techniques for forested ecosystems, none of these sources combine wide/total coverage and representation of each individual tree[5]. Reaching this level of detail is critical for dryland monitoring and management because dryland trees grow isolated and in highly variable size and density. Most current studies producing or using areal averages of tree cover, wood mass or carbon stocks in drylands are either at the very local level[12] or the information for drylands is derived from global maps[13], which are rarely trained and validated in drylands and often apply the same method on both forests and dryland vegetation[6,7,8]. Although national tree inventories exist for few dryland countries, the amount of labour required and their uncertainty are high. As a result, all existing assessments on dryland carbon stocks are highly uncertain, very difficult to validate and do not provide the means for a detailed characterization at the level of individual trees[14]. Furthermore, the contribution of different dry-mass components—wood, foliage and root mass—to the overall carbon stock is unknown at large scales.

At the same time, it remains unknown whether ecosystem models quantify the right amount of carbon and the lack of validation of global models or maps in dry areas fuels narratives of possible underestimation or overestimation of carbon stocks of drylands and their role in accelerating or mitigating climate change[12,18]. The missing information on trees at the level of individuals is decisive for improved management of woody resources in

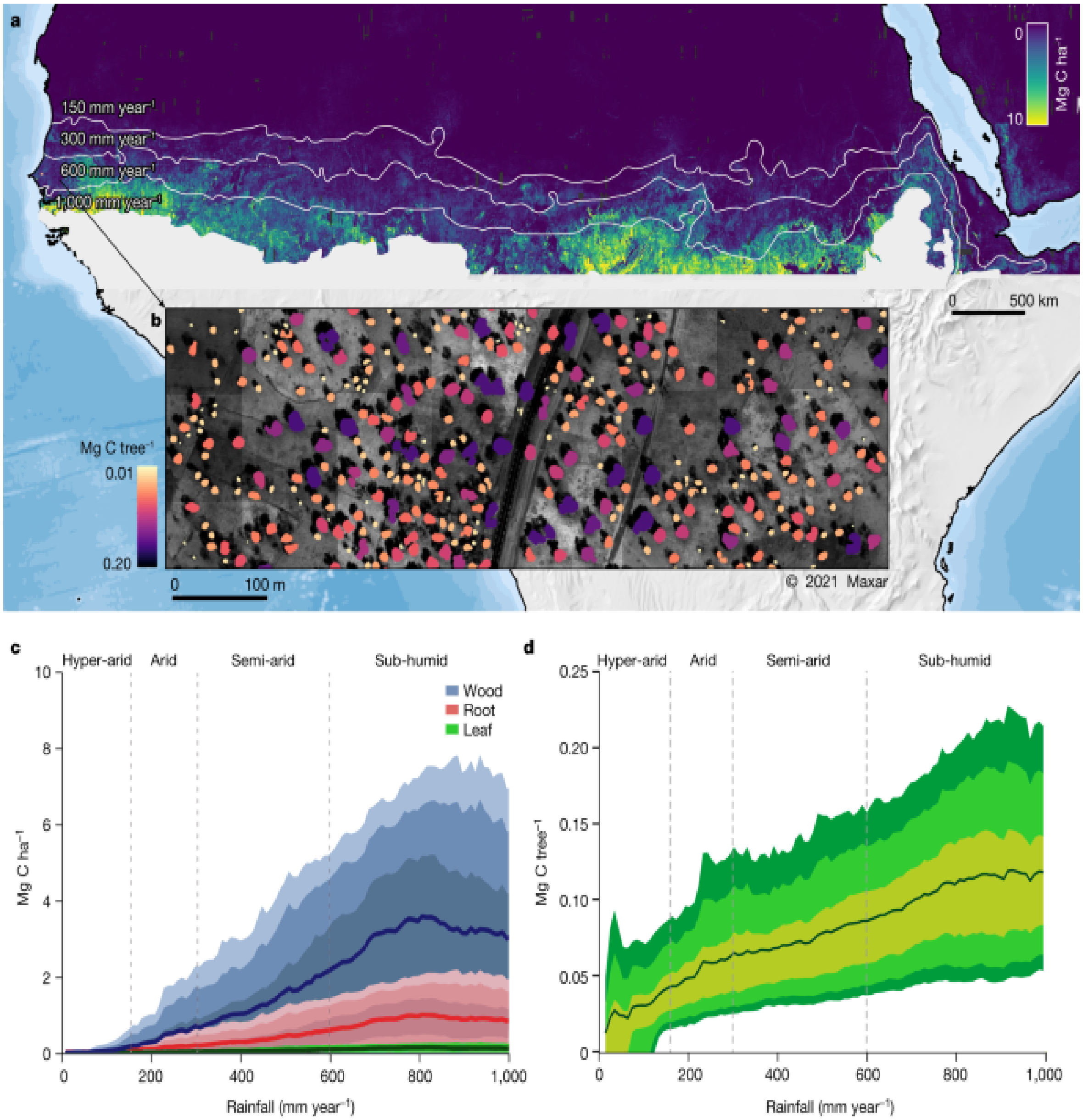
drylands: to accurately monitor deforestation spurred by clearing of trees for cropping, mining, infrastructure and urban development[24]. Furthermore, accurate monitoring of the tree resource at the level of individual trees is instrumental for tree-planting initiatives, for reporting the correct number of trees and carbon stocks for national reporting schemes, such as the Paris Agreement, or to have a reliable system that allows payments for environmental services to farmers and villages. Although deforestation and afforestation areas can be accurately mapped using current methods and data in forest ecosystems, no monitoring system exists for trees outside forests and their carbon pools[32].

At present, large amounts of funding are being allocated to dryland-restoration activities and the monitoring of success or failure is based on local inventories lacking large-scale assessments of survival rates of planted trees. The Great Green Wall of the Sahara and the Sahel initiative has recently been subject to renewed interest and increased investments[33,34,35]. This initiative was conceived to address the increasing challenges of desertification and drought, food insecurity and poverty in the wake of climate change. Yet the tracking of projects and their successfulness remains a great challenge, as no monitoring system is in place. Equally important, large-scale monitoring of single trees will create a foundation for establishing improved knowledge on the functional traits of dryland trees, such as survival, growth and mortality, controlled by the complex interplay between biotic and abiotic factors[32]. Afforestation initiatives should also be rooted in a solid ecological understanding of the local environment to avoid causing water shortages for small-holder farming systems[33].

The combined use of very-high-resolution satellite images and artificial intelligence made it possible to identify isolated trees and map their crown area at large scales, covering the western Sahara–Sahel–Sudan areas[1]. This approach of mapping individual trees has been extended to a 7.5-times-larger area covering the drylands across Africa, from the Atlantic Ocean to the Red Sea from 9.5° N to 24° N latitude between the 0 and 1,000 mm year^{-1} isohyets, using 326,523 satellite images at a 50-cm spatial resolution, and coupled with machine learning to map 9.9 billion trees (Fig. 1 and Methods). The large-scale mapping of individual tree crowns provides an unprecedented opportunity to apply allometric equations to estimate carbon

stocks derived from foliage, wood and root dry masses at local scales to large regions, here close to 10,000,000 km^2 (Extended Data Fig. 2). We take this step to assess the woody carbon pool by adding up tree-by-tree values, calculated using allometric equations to predict foliage, wood and root dry masses from crown area multiplied by the average carbon concentration (0.47). These allometric equations were established by destructive sampling of trees from 26, 27 and 5 species, respectively, selected within a rainfall gradient from 150 to 800 mm $year^{-1}$. Comparisons with allometric equations established in wetter tropical areas ensure applicability of these equations to wetter zones, at least up to 1,000 mm $year^{-1}$ rainfall[19]. We estimated the combined uncertainty from the allometric equations and the tree crown detection to be ±19.8%.

Fig. 1: Wood, foliage and root carbon of 9,947,310,221 trees with crown area >3 m^2 across 9.7 million km^2 were mapped.



a, Our study covered the southern Sahara, the Sahel and the northern Sudanian zone of Africa and showed the aggregated carbon density (foliage + wood + root) per hectare for 9,947,310,221 tree crowns from the 0–1,000 mm year^{-1} mean precipitation area. The isohyets mark the 150, 300, 600 and 1,000 mm year^{-1} rainfall zones (from north to south). **b**, Example showing the woody carbon stock of each single tree for an agroforestry area in Senegal. **c**, Mean tree carbon density at the 5th, 25th, 75th and 95th percentiles along the rainfall gradient for wood, foliage and root carbon. **d**, Mean carbon stock of individual trees at the 5th, 10th, 25th, 75th, 90th and

95th percentiles along the rainfall gradient. Our definition of a tree is a green leaf crown >3 m^2 with an associated shadow (Extended Data Fig. 1).

The information of carbon stocks of 9.9 billion trees is compared with a set of state-of-the-art TRENDY ecosystem models[23] as well as current satellite-observation-based regional carbon stock maps[6,7,8,9,10,11]. We introduce a publicly available ‘viewer’, which allows farmers, villagers, policymakers and all stakeholders to retrieve the foliage wood and root masses and the corresponding carbon stock of each tree using a mobile device. We expect that this could improve not only the amount of information available but also the reporting and monitoring of trees and their carbon stocks at various scales, from the individual field plot to country scales.

Carbon stocks at the tree level

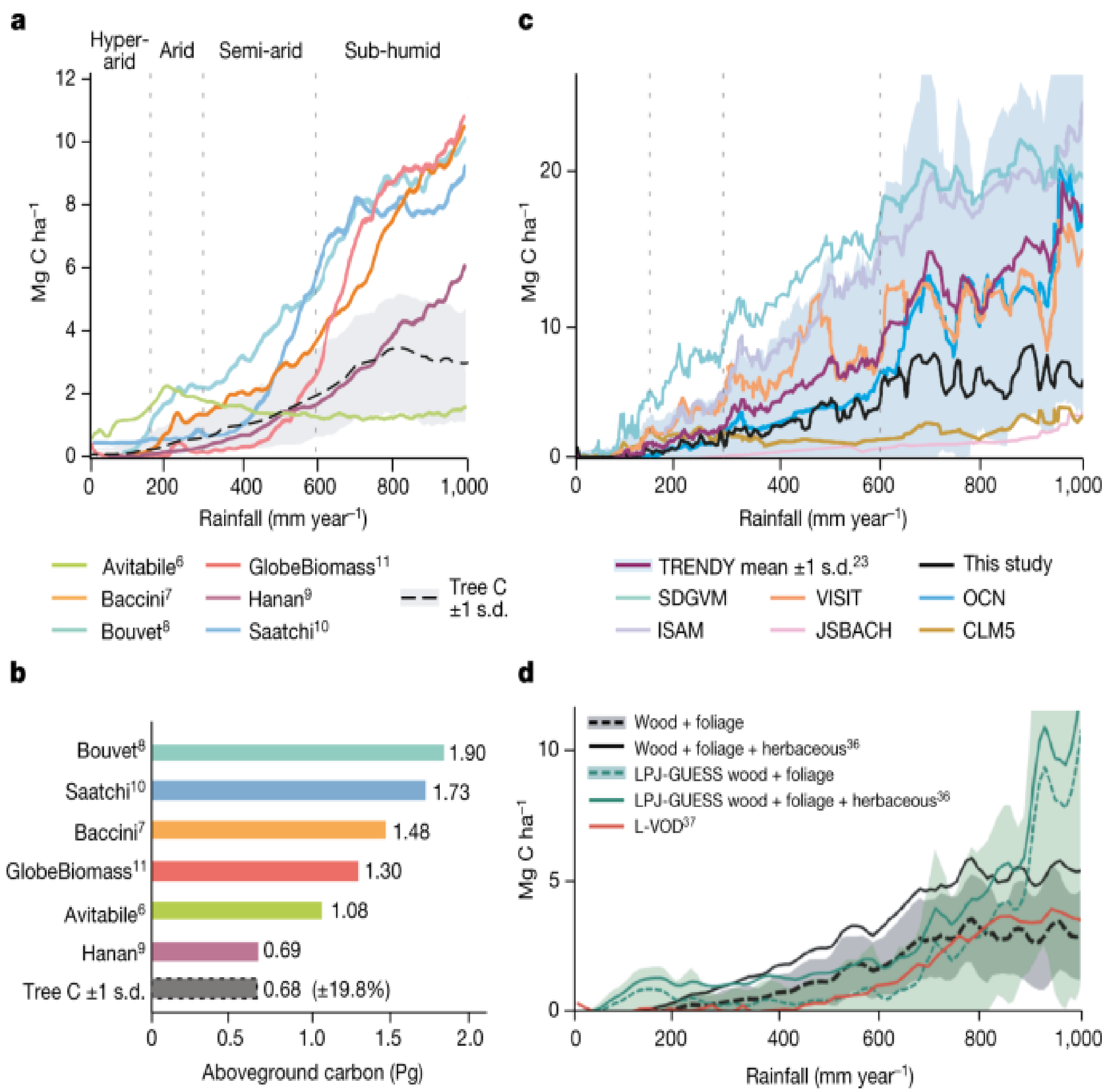
We applied a deep-learning-based tree mapping on a large number of satellite images and measured 9,947,310,221 tree crowns: all woody plants with a shadow and a crown area >3 m^2 from the hyper-arid (0–150 mm $year^{-1}$), arid (150–300 mm $year^{-1}$), semi-arid (300–600 mm $year^{-1}$) and dry sub-humid (600–1,000 mm $year^{-1}$) rainfall zones of tropical Africa north of the Equator and south of the Sahara (Fig. 1). The average carbon stock of a single tree is 51 kg C in the hyper-arid, 63 kg C in the arid, 72 kg C in the semi-arid and 98 kg C in the sub-humid zone. The individual tree information was projected to the area by calculating the carbon density in Mg C ha^{-1}, which was on average 0.03 Mg C ha^{-1} in the hyper-arid, 0.54 Mg C ha^{-1} in the arid, 1.54 Mg C ha^{-1} in the semi-arid and 3.73 Mg C ha^{-1} in the sub-humid zone. Although foliage mass has a small overall fraction of the total dry mass (3%), it is an important variable for quantification of browse potential and serves as a proxy for other ecosystem processes, such as transpiration, photosynthesis and nutrient cycling. The proportion of root mass is on average 15–20% of the total mass.

Current carbon map and model comparisons

We compared our aboveground carbon-density maps (foliage + wood) derived from individual trees with current state-of-the-art maps (Fig. 2 and

Extended Data Fig. 3) available at moderate spatial resolutions of 30–1,000 m. The temporal dynamics were assessed by low-frequency passive microwaves (L-VOD)[36,37], which has emerged as a tool for the assessment of carbon stock dynamics at the 25 × 25-km spatial scale (Extended Data Fig. 4). Moreover, we compared carbon-density maps and dynamics with dynamic ecosystem models from the TRENDY database with a 50 × 50-km grid cell size[23]. None of these maps were designed specifically for drylands; most dynamic ecosystem models and satellite-based models are developed and trained for forest ecosystems and, in the case of the TRENDY models, used meteorological forcings and prescribed vegetation maps that contain further uncertainties for comparative purposes.

Fig. 2: Comparisons between current aboveground carbon-density maps and our estimations derived from 9.9 billion trees.



a, Aboveground carbon density from state-of-the-art maps using satellite data[6,7,8,9,10,11]. Tree carbon from this study is derived from wood + foliage carbon plotted with ±1 standard deviation in the grey zone. **b**, Aboveground carbon stocks aggregated over the 0–1,000 mm year^{-1} rainfall zone. Our estimations (grey colour) of 0.68 Pg are wood + foliage carbon. The combined uncertainty from neural net area mapping, tree crown omission and commission errors, and allometric conversion of tree crowns into tree wood, foliage and root carbon was ±19.8% (Methods). **c**, Vegetation carbon density from the mean of 14 TRENDY dynamic ecosystem models and data from six individual TRENDY models for aboveground and belowground carbon[23] are compared with our tree carbon with aboveground herbaceous

carbon added from passive microwaves[36]. **d**, Aboveground carbon density from the LPJ-GUESS model[23], selected here as it uses trees outside the prescribed forest fraction, and our estimations are compared along the rainfall gradient. L-VOD[37] was converted to carbon density using coefficients from a linear correlation with our map (Extended Data Fig. 4). Aboveground herbaceous carbon was derived from ref. [36]. The sample size for the 0–1,000 mm $year^{-1}$ rainfall zone was 9,947,310,221 tree crowns >3 m^2.

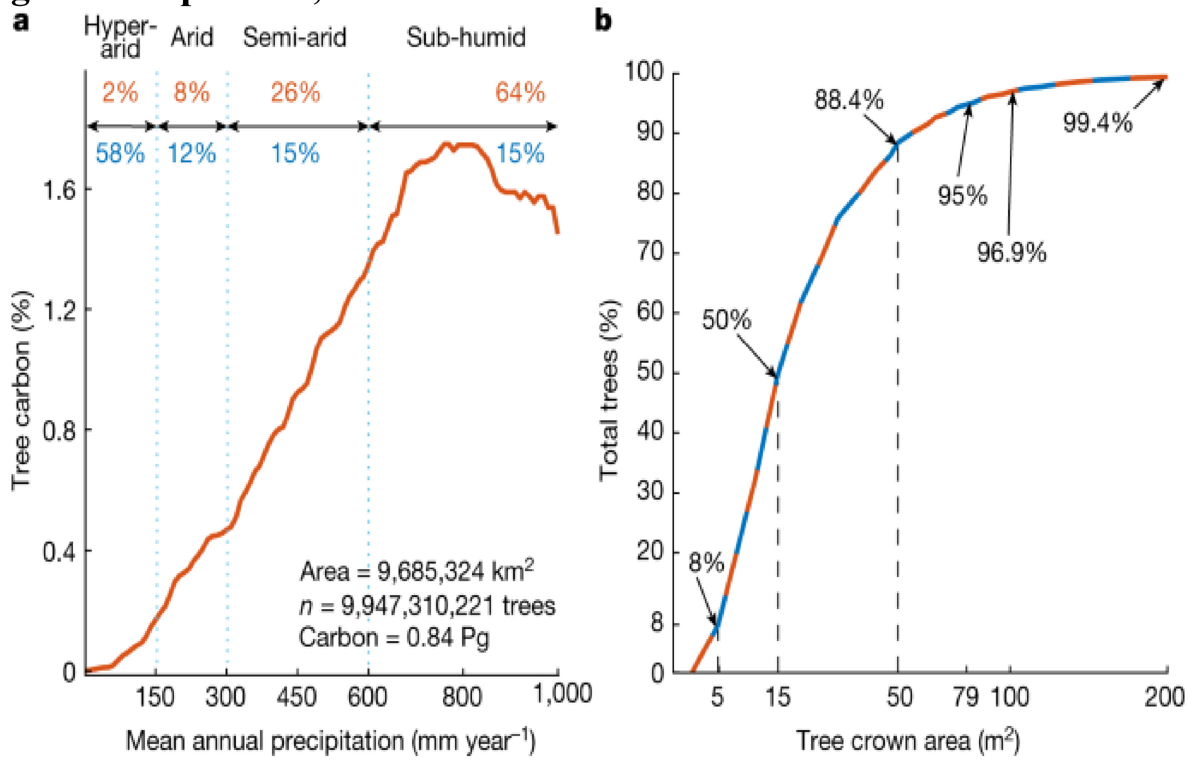
Existing carbon-density maps compare differently with our assessment based on individual trees and there is little spatial agreement among the maps (Fig. 2a,b). Notably, although areas of scattered trees having a relatively low carbon density are largely mapped as zero carbon in previous maps except for ref. [9], areas of denser tree cover and some areas typically without trees, such as wetlands, irrigated croplands and desert mountains, have considerably higher values than our assessment. This leads to an overall higher carbon stock of the area compared with our results. Although we do not map herbaceous vegetation in our study, the tree cover we map can be used to disaggregate herbaceous vegetation from trees (Extended Data Fig. 5).

At regional scales, dynamic ecosystem model vegetation carbon shows a considerable variability, but the mean follows our estimates of herbaceous, wood, foliage and root carbon along the rainfall gradient (Fig. 2c). Notably, whereas previous studies assumed that ecosystem models underestimated dryland carbon stocks, our results show overall higher values from the model outputs as compared with the assessment based on individual trees, although large variations between models exist. Only considering aboveground carbon, the example of LPJ-GUESS shows slightly lower values than our assessment up to about 800 mm $year^{-1}$ rainfall (Fig. 2d).

Both ecosystem models and previous satellite-based carbon maps diverge markedly from our results beyond 700–800 mm $year^{-1}$ rainfall. All other maps assume a continuous increase beyond this rainfall zone, yet our results reach a plateau at 800 mm $year^{-1}$ and no further increase in carbon is observed with higher rainfall up to 1,000 mm $year^{-1}$. We acknowledge that the uncertainty of our results increases with denser canopy cover and that we

miss all understory vegetation. However, statistical evaluations of the rainfall–tree density relationship from our data indicate that neither carbon stocks per tree (Fig. 1d) nor tree cover further increased between 800 and 1,000 mm year^{-1} rainfall (Fig. 3a). Trees with crown area <50 m^2 make up 88% of the total number of trees, whereas trees in the semi-arid and sub-humid zones constitute 90% of the total carbon in our study (Fig. 3).

Fig. 3: Precipitation, tree carbon and crown area.



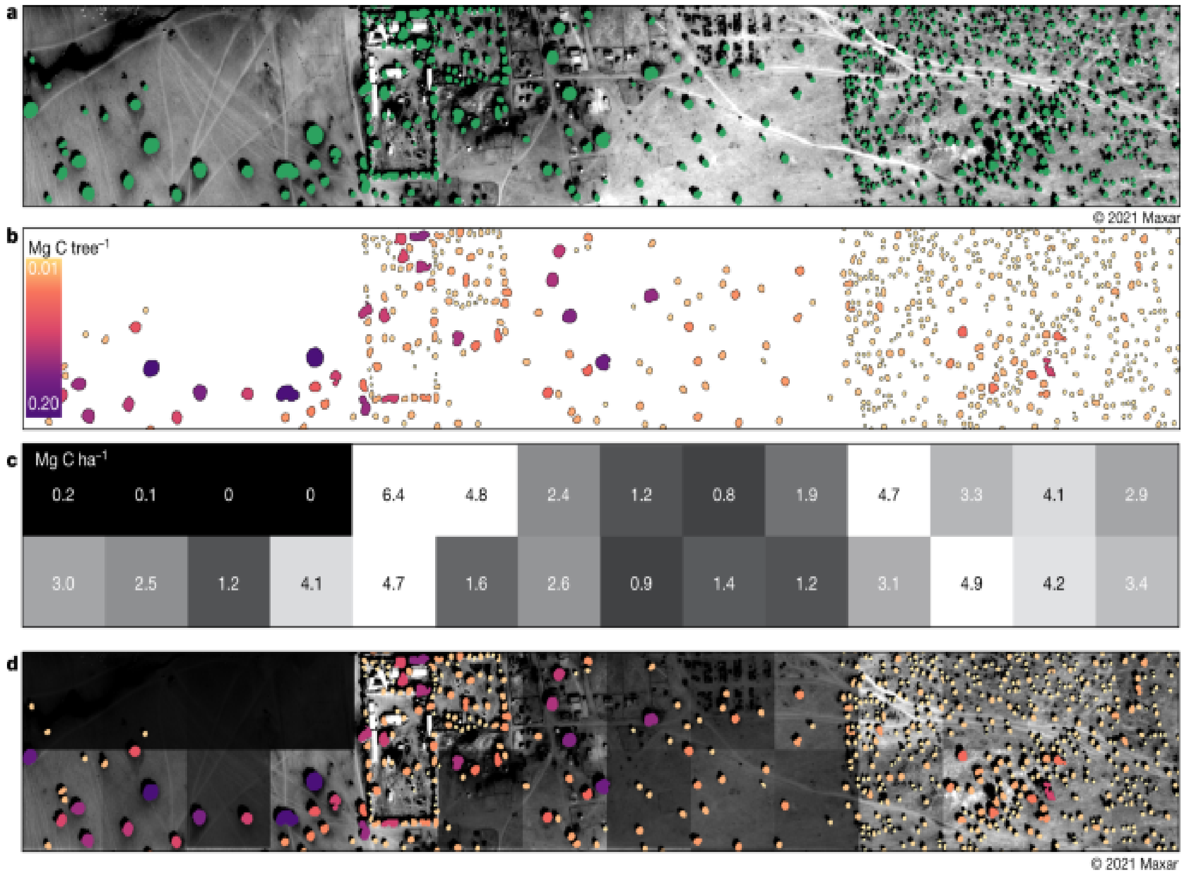
a, The tree carbon probability density function computed along the rainfall gradient of the study from the hyper-arid (0–150 mm year^{-1}), arid (150–300 mm year^{-1}), semi-arid (300–600 mm year^{-1}) and dry sub-humid (600–1,000 mm year^{-1}) rainfall zones. The percentage area of each semi-arid zone is shown in blue and the percentage of total carbon in red. The increasing tree carbon probability function shows the importance of precipitation for tree carbon in semi-arid regions. Most tree carbon is found in the semi-arid (26%) and dry sub-humid zones (64%), which represent only 30% of the area in our study. The per cent carbon density contribution by rainfall zones is linearly related to the tree carbon density (Mg C ha^{-1}) reported in Fig. 1c by a factor of 2.5. **b**, A total of 88.4% of our mapped trees had crown areas <50 m^2. The average tree crown area in the 0–150 mm year^{-1} zone was 15.1

m^2, for the 150–300 mm $year^{-1}$ zone it was 18.4 m^2, for the 300–600 mm $year^{-1}$ zone it was 20.9 m^2 and for the 600–1,000 mm $year^{-1}$ zone it was 28.1 m^2. Only 11.6% of our mapped trees had crown areas >50 m^2 and less than 0.6% had crown areas >200 m^2.

Application at the tree level

The comparison with dynamic global vegetation models and existing biomass maps shows some similar patterns at coarse scale, yet none of these maps can be used to derive information at the level of individual trees needed to support policymakers and decision-makers. For this reason, we introduce a viewer (Fig. 4), which is built on Mapbox and OpenStreetMap, and can be accessed online by everyone and from anywhere. The viewer includes all 9.9 billion trees as objects, and the wood, foliage and root mass can be accessed individually for each of them. As an example, we show the area of Widou Thiengoly, an area in Senegal in which tree planting for the Great Green Wall has been promoted over the past decades (Fig. 4a). Although previous assessments on the success of tree plantations were based on narratives, visual interpretations or site visits, the viewer provides an unbiased tool for evaluating success and failure of initiatives, as well as quantifying the carbon stocks gained by each planted tree or lost by each removed/dead tree. The example shown in Fig. 4 illustrates that high-density plantations in this arid region reach carbon-density values of about 5 Mg C ha^{-1} (Fig. 1c), but the survival rate of planted trees has been a long-lasting concern that needs to be carefully monitored to be able to assess the efficacy of Sahelian tree-planting programmes.

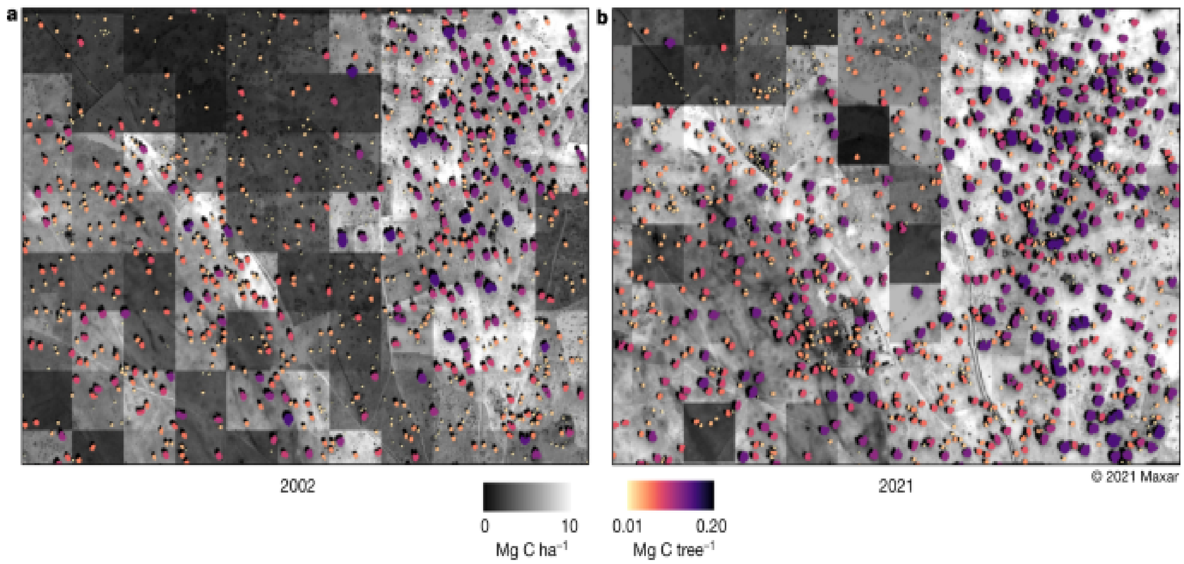
Fig. 4: Different components of the viewer.



This example shows Widou Thiengoly in semi-arid Senegal surrounded by tree plantations, which are partly related to a Great Green Wall[34] project aiming to increase tree density and improve livelihoods in the Sahel. **a**, Tree crown segmentations from the neural net mapping. **b**, Wood, foliage and root carbon calculated for each tree (Methods). **c**, Carbon density per hectare aggregated from carbon stocks of single trees to the hectare scale. **d**, Our viewer includes all information from **a** to **c**. This online tool provides information on crown area; foliage, wood and root carbon of single trees; and aggregates carbon to the hectare scale. These data can be accessed by policymakers and stakeholders to monitor areas of interest. The viewer can be accessed at https://trees.pgc.umn.edu/app.

Another example shows an agroforestry region in Senegal, north of Khombole that has a relatively high density of trees, which has increased the carbon stocks of the region considerably. The example area shown in Fig. 5 has almost doubled carbon density between 2002 and 2021 (Fig. 5).

Fig. 5: Monitoring at the level of single trees from Khombole, Senegal.



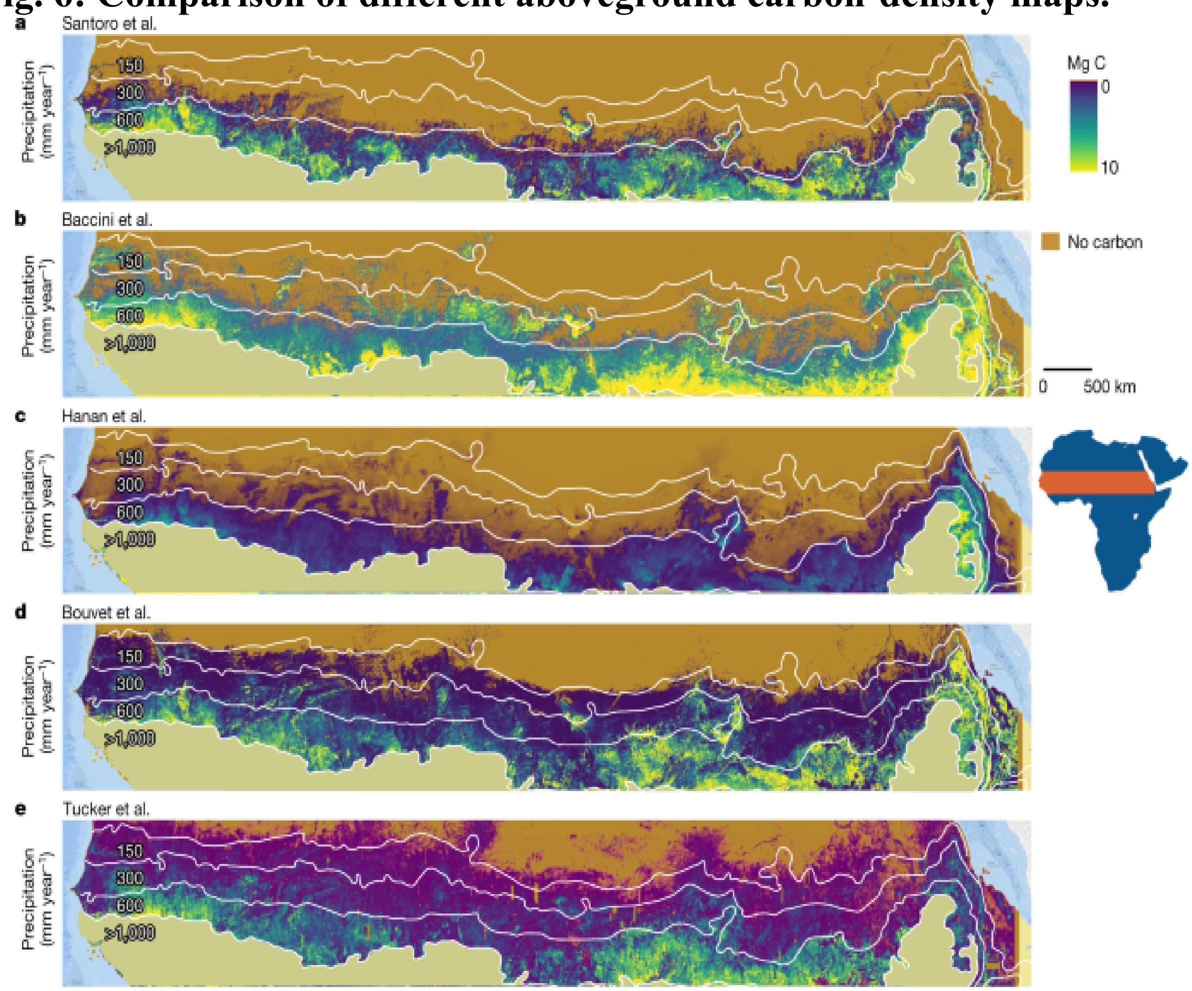
a,**b**, A 50-cm-scale image from 2002 (**a**) and a 50-cm-scale satellite image from 2021 (**b**) showing an agroforestry area at the same location. Tree cover has increased between 2002 and 2021 and the average carbon density of both areas was calculated and increased from 6 to 10 Mg ha^{-1}. A large number of trees grow on farmlands, keeping the soils fertile and reducing the need for fallow periods. The greyscale of the background images indicates the carbon density per hectare, whereas the colour scale shows the carbon content of individual trees. This is a good example of the tree restoration monitoring potential in our study area.

Discussion

Our assessment is a large-scale estimation of wood, foliage and root carbon at the level of individual trees. The finding that global ecosystem models and previous carbon-density maps estimate higher carbon stocks in African drylands compared with our assessment based on 9.9 billion individual trees seems surprising, as current tree-cover maps are not able to correctly account for scattered trees and thus should considerably underestimate the number of trees in these areas[1]. The explanation for this apparent paradox—higher tree cover but less carbon—is related to the fact that previous models are rarely developed, trained and validated with plots of very sparse tree cover, thus leaving high uncertainty for drylands with scattered trees. Consequently, areas with scattered trees are often represented by zero values (Fig. 6), whereas the carbon density of larger groups of trees may be

overestimated in previous assessments, as these areas are wrongly considered as dense forests. In essence, most previous assessments do not accurately map carbon density below 10 Mg C ha^{-1}, if at all, and may overestimate the carbon stocks of dryland 'forests'. Moreover, if the region is taken as a whole, green crops and herbaceous vegetation affect optical images, whereas steep topography and wetlands/irrigated areas affect the radar backscatter, both predicting higher carbon stocks than our estimations. Although we used allometric equations specifically developed from locally sampled field data[19], 95% of the trees we mapped had a crown area <78 m^2. This introduces a small uncertainty in carbon values for the 5% of tree crowns >78 m^2 in more humid areas, where trees are taller and/or larger.

Fig. 6: Comparison of different aboveground carbon-density maps.



We show the aboveground carbon density for our study area compared with different sources. Areas beyond 1,000 mm $year^{-1}$ rainfall are masked out.

Data are from Santoro et al.[11] (**a**), Baccini et al.[7] (**b**), Hanan et al.[9] (**c**), Bouvet et al.[8] (**d**) and Tucker et al. (this paper; foliage + wood) (**e**). See also Fig. 2a and Extended Data Fig. 3.

Nevertheless, the divergence between our results and previous assessments in higher-rainfall zones needs to be further investigated and our maps should be used with caution beyond 800 mm year^{-1} rainfall. The indirect inclusion of the tree height and the application of the same equation to all tree species are uncertainty factors that will be assessed in future versions of the dataset. Finally, the fact that larger trees shade out smaller trees in areas of dense tree cover makes the method based on individual tree counting less suited to more humid areas.

Herbaceous dry mass can contribute considerably to the annual carbon density. However, most herbaceous plants of the region are annuals that die off each year and do not constitute a residual carbon stock but have a high inter-annual variability. The herbaceous mass used in our study[36] shows the seasonal peak value, which drops by about 25% within only a few weeks (Extended Data Figs. 1a and 5). Traditionally, remotely sensed separation of herbaceous vegetation from woody foliage is challenging with both optical and radar satellite data. We overcome this by measuring individual tree crown areas.

The carbon difference between ecosystem models and our study can be explained by different forest fractions assumed by each model (Extended Data Fig. 6). Most of the dynamic global vegetation models do not simulate trees outside forests and woody carbon is usually a sum of predefined forest areas. Differences may also result from a simplistic implementation of disturbances, in particular, fire, grazing and the fact that we did not include belowground herbaceous carbon in our estimates. Still, the results of the dynamic vegetation models are closer to our estimations than previously assumed and the inclusion of our data may improve future modelling results, leading to more realistic forecasts of the impact of climate change on drylands.

Dryland trees are not only a carbon stock but also provide ecosystem services that are valuable to the environment and support local livelihoods, including timber, fuel wood, protection against soil erosion and loss, soil

fertilization, shade and nutrition for tree crops[15]. The benefits of increased tree cover are many and establishing an operational monitoring system for dryland trees is critically needed. The dynamics of growth and mortality of trees outside forests goes undetected by conventional monitoring systems based on satellite imagery with a spatial resolution >10 m. Although our current assessment at the level of individual trees does not yet include a temporal dimension (except for the exemplary case provided in Fig. 5), it is a baseline of the number, mass and carbon stock of trees outside forests at the sub-continental scale. The publicly available viewer makes this information accessible for scientists, policymakers, stakeholders and individual farmers, who can easily quantify woody carbon stocks of a given area, down to the level of a single tree growing in a private yard.

A next step will be adding a temporal dimension to the wall-to-wall mapping we describe and we expect it to be possible from this source of data, at least with decadal time steps. This will facilitate addressing the impact of droughts, restoration and policies at various scales, down to the level of individual trees. High spatial resolution is the key to improved tree inventories in drylands. The ever-increasing availability of satellite images will make continental-scale assessments of carbon pools and dynamics at the individual tree level realistic in near-real time. This will be key to developing robust schemes for dryland management plans needed to achieve the United Nations' Sustainable Development Goals. Our paper is a step in that process.

Methods

Overview

This study establishes a framework for mapping carbon stocks at the level of individual trees at a sub-continental scale in semi-arid sub-Saharan Africa north of the Equator. We used satellite imagery from the early dry season (Extended Data Fig. 1). The deep learning method developed by a previous study[1] allowed us to map billions of discrete tree crowns at the 50-cm scale from West Africa to the Red Sea. Then we used allometry to convert tree crown area into tree wood, foliage and root carbon for the 0–1,000 mm year^{-1} precipitation zone in which our allometry was collected (Extended

Data Fig. 2). We introduce a viewer that enables the billions of trees to be viewed at different scales, with information on location, metadata of the Maxar satellite image used, tree crown area and the estimated wood, foliage and root carbon content based on our allometry (Fig. 4). We also make available our output data for the 1,000 mm year^{-1} precipitation zone southward to 9.5° N latitude with information on location, precipitation, metadata of the Maxar satellite image used, tree crown area, tree wood carbon, tree root carbon and tree leaf carbon.

Satellite imagery

We used 326,523 Maxar multispectral images from the QuickBird-2, GeoEye-1, WorldView-2 and WorldView-3 satellites collected from 2002 to 2020 from November to March from 9.5° N to 24° N latitude within Universal Transverse Mercator (UTM) zones 28–37 for Africa (Extended Data Table 1a). These images were obtained by NASA through the NextView License from the National Geospatial-Intelligence Agency. Data were assembled over several years with a focus on later years to achieve a relatively recent and complete wall-to-wall coverage.

When using satellite data from different satellites over several years, with varying sun–target–satellite angles, with varying radiometric calibration of satellite spectral bands and different atmospheric compositions through which the surface is imaged, there are two possibilities for using hundreds of thousands of satellite images together quantitatively. One approach, used extensively in NASA's, NOAA's and the European Space Agency's Earth-viewing satellite programmes, is to quantitatively inter-calibrate radiometrically the satellite channels through time; correct these data for time-dependent atmospheric effects such as aerosols, clouds, haze, smoke, dust and other atmospheric constituent effects and then normalize the viewing perspective to the same sun–target–satellite angle[38]. Another approach is to use the satellite data as collected; assemble training data of trees viewed from different satellites under different sun–target–satellite angles, different times, different atmospheric conditions and use machine learning with high-performance computing to perform the tree mapping at the 50-cm scale. The key to successful machine learning is to account for all the sources of variation within the domain of study in the training data to

ensure accurate identification of trees under all circumstances. We included trees viewed substantially off-nadir, trees collected under different aerosol optical thicknesses, trees collected under cirrus cloud conditions, trees viewed in the forward and backward scan directions, trees on sandy soils, trees on clay soils, trees on burn scars, trees in laterite areas and trees in riverine settings. Our training data were collected by one team member and are a carefully selected manual delineation of 89,899 individual trees under a range of atmospheric conditions, viewing perspectives and ecological settings.

All multispectral and panchromatic bands associated with our Maxar images were orthorectified to a common mapping basis. We next pan-sharpened all multispectral bands to the 0.5-m scale with the associated panchromatic band. The absolute locational uncertainty of pixels at the 0.5-m scale from orbit is approximately ±11 m, considering the root-mean-square location errors among the QuickBird-2, GeoEye-1, WorldView-2 and WorldView-3 satellites (Extended Data Table 1). We formed the normalized difference vegetation index (NDVI)[39] from every image in the traditional way from the pan-sharpened red and near-infrared bands. We also associated the panchromatic band with the NDVI band and ensured that the panchromatic and NDVI bands were highly co-registered. The NDVI was used to distinguish tree crowns from non-vegetated background because the images were taken from a period when only woody plants were photosynthetically active in this area[36]. Our training data were labelled on images from the early dry season when only trees have green leaves. Because most semi-arid savannah trees continue to photosynthesize in the early dry season after herbaceous vegetation senesces, green leaf tree crowns are easily mapped because of their higher NDVI values than their senescent herbaceous vegetation surroundings. We substantiate this by analysis of 308 individual trees using NDVI time series with 4-m PlanetScope imagery that emphasized the importance of satellite data from the November, December and January early dry-season months (Extended Data Fig. 1).

We next formed our data into mosaics by applying a set of decision rules, resulting in a collection of 16 × 16-km tiles within each UTM zone from 9.5° N to 24° N latitude for Africa. The initial round of scoring considered percentage cloud cover, sun elevation angle and sensor off-nadir angle: preference was given to imagery that had lower cloud cover, then higher sun

elevation angle and finally view angles closest to nadir. In the second round of scoring, selections were assigned priority to favour early dry-season months and off-nadir view angles: preference was given to imagery from November, December and January with off-nadir angle less than ±15°; second to imagery from November to January with off-nadir angle between ±15° and ±30°; third to imagery from February or March with off-nadir angle less than ±15°; and finally to imagery from February or March with off-nadir angle between ±15° and ±30°. Image mosaics were necessary to eliminate multiple counting of trees. We formed mosaics using 94,502 images for tree segmentation, with 94% of these being from November, December and January. Ninety percent of our selected mosaic imagery was within ±15° of nadir, 87% were acquired between 2010 and 2020 and 94% were from the early dry season (Extended Data Fig. 7). A summary of month, year, solar elevation and off-nadir angle by UTM zone can be found in Supplemental Information Fig. 1.

Possible obscuration of the surface by clouds totalled 4.1% of our input mosaic data area and aerosol optical depth >0.6 at 470-nm (ref. [40]) areas totalled 3.4% of our input data. However, we mapped 691,477,772 trees in our possible cloud-cover-affected and aerosol-affected areas, indicating that cloud and aerosol effects were lower than these numbers. In addition, 0.9% of our input data did not process. We include a data layer in our viewer for these three conditions.

Mapping tree crowns with deep learning

We used convolutional neural network models developed by a previous study[1]. The models were trained with manually delineated and annotated 89,899 individual trees along a north–south gradient from 0 to 1,000 mm year^{-1} rainfall[1]. Only features that showed a distinct crown area and associated shadow were included, which excluded small bushes, grass tussocks, rocks and other features that might have green leaves or cast a shadow from our classification. All training data and model training was done in UTM zones 28 and 29. Because tree floristic diversity in the 0–1,000 mm year^{-1} zone of our study is highly similar from the Atlantic Ocean to the Red Sea across Africa[41,42,43], we added no further training data as our study moved further eastward. We used state-of-the-art deep learning to

segment trees crowns at the 50-cm scale[1]. We used two different models based on a U-Net architecture, one for lower-rainfall desert regions with <150 mm $year^{-1}$ precipitation and one for regions with average annual precipitation >150 mm $year^{-1}$. Details about the network architecture, training process and hyperparameter choices can be found in ref. [1]. Previous evaluation showed that early dry-season images performed better than late dry-season images, which was a limitation of our previous study. We reduced this error by using early dry-season images with only 6% of our area being covered by images from February and March. The models were also designed to separate clumped trees by highlighting spaces between different crowns during the learning process, similar to a strategy for separating touching cells in microscopic imagery[22].

Allometry

Very-high-resolution satellite images and deep learning have achieved mapping of individual trees over large areas[1]. Each tree is georeferenced in the satellite data and defined by crown area. The challenge was to develop allometric equations for foliage, wood and root dry masses or carbon based on crown area regardless of species. This was met by reanalysing existing Sahelian and Sudanian woody plant data from destructive sampling. Overall, the seasonal maximum foliage, wood and root dry masses were measured on 900, 698 and 26 trees or shrubs from 27, 26 and 5 species, respectively, for which crown area was also measured. Several allometric regression models tested for foliage, wood or root masses are power functions and independent of species. All the regression outputs were inter-compared for fit indicators, by systematic estimates of prediction uncertainty and by root-to-wood ratios and foliage-to-wood ratios over the range of crown areas. This resulted in a set of ordinary least squares log–log equations with crown area as the independent variable. The Sahelian and Sudanian allometry equations were also compared with published allometry equations for tropical trees, primarily from more humid tropics, which are generally based on stem diameter, tree height and wood density. Our allometric predictions are within the range of other allometry predictions, reinforcing the confidence in their use beyond the Sahelian and Sudanian domains into sub-humid savannahs for discrete trees[19].

On the basis of ref. [19], we predicted the wood (w), foliage (f) and root (r) dry mass as functions of the crown area (A) of a single tree as:

$$\begin{array}{c}{\text{mass}}_{{\rm{w}}}(A)=3.9448\times {A}^{1.1068}\,({N}_{{\rm{w}}}=698)\\ {\text{mass}}_{{\rm{f}}}(A)=0.2693\times {A}^{0.9441}\,({N}_{{\rm{f}}}=900)\\ {\text{mass}}_{{\rm{r}}}(A)=0.8339\times {A}^{1.1730}\,({N}_{{\rm{r}}}=26)\end{array}$$

The tree mass components of wood, leaves and roots were combined to predict the total mass(A) in kg of a tree from its crown area A in m^2:

$$\text{mass}\left(A\right)={\text{mass}}_{{\rm{w}}}\left(A\right)+{\text{mass}}_{{\rm{f}}}\left(A\right)+{\text{mass}}_{{\rm{r}}}\left(A\right)$$

As in ref. [1], a crown area of size $A > 200\ m^2$ was split into $\lfloor A/100 \rfloor$ areas of size $100\ m^2$ and one area with the remaining m^2 if necessary. We converted dry mass to carbon by multiplying with a factor of 0.47 (ref. [44]).

Uncertainty analysis

We evaluated the uncertainty of our tree crown area mapping and carbon estimation in two ways. First, we quantified our tree crown mapping omission and commission errors by inspecting randomly selected areas from UTM zones 28–37, validating that our neural network generalized over UTM zones consistently (Extended Data Fig. 8).

Second, we quantified the relative error of our tree crown area estimation. We consider the uncertainty Δ_x of a quantity x and the corresponding relative uncertainty δ_x defined by the absolute and relative error, respectively[45]. To assess the relative error in crown area estimation resulting from errors by the neural network, we considered external validation data from ref. [1], which were not used in the model-building process. We considered expert-labelled tree crowns as well as the predicted tree crowns from 78 plots of 256 × 256 pixels. The hand-labelled set contained 5,925 trees and the system delineated

5,915 trees. The total hand-labelled tree crown area was 118,327 m^2 and the neural network predicted 121,898 m^2. This gave a relative error in crown area mapping of $\delta_{area} = 3.3\%$. We matched expert-labelled and predicted tree crowns and computed the root-mean-square error (RMSE) per tree, taking overlapping areas and missed trees into account (see Extended Data Fig. 8). We estimated the allometric uncertainty ($\delta_{allometric}$) using the data from ref. [19] (see below). The two relative errors δ_{area} and $\delta_{allometric}$ were combined to an overall uncertainty estimate for the carbon prediction of ±19.8% (see below).

Omission and commission errors

We evaluated our tree crown mapping accuracy by analysis of 1,028 randomly selected 512 × 256-pixel areas over the 9.5° N to 24° N latitude within UTM zones 28–37. Because the drier 60% of our study area only contains 1% of the 9,947,310,221 trees we mapped in the 0–1,000 mm $year^{-1}$ rainfall zone, we applied an 80% bias for selecting evaluation areas above the 200 mm $year^{-1}$ precipitation line[46], as >98% of tree identifications were above the 200 mm $year^{-1}$ precipitation isoline. Identified tree polygons were further categorized into tree crown area classes from 0–15 m^2, 15–50 m^2, 50–200 m^2 and >200 m^2, with a total of 50,570 trees evaluated. Although a previous study reported greatest uncertainty in both the smallest and largest area classes[1], our more expansive work found the greatest uncertainty in our smallest tree class. We excluded from evaluation any tiles that had annual precipitation[46] >1,000 mm $year^{-1}$ and all areas that were devoid of vegetation, leaving us with 850 areas.

Seven members of our team evaluated the accuracy in terms of commission and omission by tree crown area classes for the 850 areas. Input data provided for every area were the NDVI layer, the panchromatic shadow layer and the neural net mapping results in each of the four crown area classes. Ancillary data available to evaluators included the centre coordinates for comparison with Google Earth data, the Funk et al.[46] rainfall, the acquisition date of the area evaluated and the viewing perspective.

We identified areas wrongly classified as tree crowns (commission errors), missed trees (omission errors) and crown areas corresponding to clumped trees (Extended Data Fig. 8). Clumped trees were most common for >200 m^2 tree crown area. They were rare in the 3–15 m^2 and 15–50 m^2 tree classes, which comprise 88% of our tree crowns. In the 850 patches, the number of trees ranged from one tree to 326 trees, with a total of 50,570 trees evaluated and 3,765 errors identified. Overall, the commission and omission error rates were 4.9% and 2.7%, respectively, a net uncertainty of 2.2%.

Allometric uncertainty estimation

The prediction of tree carbon from the crown area for a single tree based on crown area alone is inherently uncertain[47,48]. As the allometric equations are based on three different datasets, we compute their uncertainties independently, combine them and put them in relation to the total carbon measured in the three datasets.

The allometric equations were established using an optimal least-squares fit of an affine linear model predicting the logarithmic carbon from the logarithmic tree crown area[19]. To estimate the uncertainty of the allometric equations, we repeated the fitting using random subsampling. The datasets were randomly split into training data (80%) for fitting the allometric equations and validation data (20%) for assessing the uncertainty. For example, from the root measurements, $(({A}_{1},{y}_{1}),\ldots ,({A}_{{N}_{{\rm{r}}}}\,,{y}_{{N}_{{\rm{r}}}}))$, we compute ${\mu }_{{\rm{r}}}=\frac{1}{{N}_{{\rm{r}}}}\mathop{\sum }\limits_{i=1}^{{N}_{{\rm{r}}}}{y}_{i}$ and ${\hat{\mu }}_{{\rm{r}}}=\frac{1}{{N}_{{\rm{r}}}}\mathop{\sum }\limits_{i=1}^{{N}_{{\rm{r}}}}{\text{mass}}_{{\rm{r}}}({A}_{i})$. The corresponding error is ${\varDelta }_{{\rm{r}}}=|{\mu }_{{\rm{r}}}-{\hat{\mu }}_{{\rm{r}}}|$.

Because the total carbon for a tree with a certain crown area is the sum of the three carbon components, we add the absolute uncertainties assuming independence[45].

$${\varDelta }_{{\rm{a}}{\rm{l}}{\rm{l}}{\rm{o}}{\rm{m}}{\rm{e}}{\rm{t}}{\rm{r}}{\rm{i}}{\rm{c}}}\simeq \sqrt{{\varDelta }_{{\rm{f}}}^{2}+{\varDelta }_{{\rm{w}}}^{2}+{\varDelta }_{{\rm{r}}}^{2}}$$

and compute the relative uncertainty as $\,{\delta }_{\text{allometric}}=\frac{{\varDelta }_{\text{allometric}}}{{\mu }_{\text{mass}}}$, in which the average mass μ_{mass} is given by the sum of the averages for wood (μ_{w}), leaves (μ_{f}) and root (μ_{r}). This process was repeated ten times, resulting in a mean relative uncertainty of

$${\bar{\delta }}_{{\rm{allometric}}}=19.5 \% .$$

Total carbon uncertainty

We combine the uncertainties from the neural net mapping and our allometric equations, which can be viewed as considering $(1 + A)\cdot(1 + B)$ with A and B being random variables with standard deviations δ_{area} and $\delta_{\mathrm{allometric}}$. Neglecting higher-order and interaction terms, we combine the two sources of uncertainty to $\delta \simeq \sqrt{{\delta }_{{\rm{area}}}^{2}+{\bar{\delta }}_{{\rm{allometric}}}^{2}}$, resulting in an uncertainty in total tree carbon for our study of ±19.8%. See also Extended Data Fig. 9 for the RMSEs of our predicted crown areas calculated on external validation data from ref. [1], binned on the basis of the 50th quantiles of the hand-labelled crown areas and converted also into carbon. Extended Data Fig. 10 is a flow diagram summarizing our methods.

Our viewer

Visualizing our large tree-mapping dataset in an interactive format was essential for quality-control purposes, exploration of the data and hypothesis creation. Creating a web-based viewer serves the purpose of being the initial point of interaction with our dataset for fellow researchers, local stakeholders or the general public. The visualization of more than 10 billion trees in a web browser required maintaining performance, interactivity and individual metadata for each polygon. Users should be able to zoom in to any area within the dataset to view individual tree polygons and query their

statistics while at the same time accurately depicting the overall trends of the dataset at lower zoom levels. The visualization also needed to clearly denote where data were missing or possibly affected by clouds or aerosols. Finally, the extent and origin of the source imagery, its acquisition date and a preview of the imagery needed to be available. To accomplish these goals, a vector-tile-based approach was taken, with the data visualized in a Mapbox GL JS map within a React web application. To create vector tiles covering the entire study area, we developed a data-processing pipeline using high-performance computing resources to transform the data into compatible formats, as well as to package, optimize and combine the vector tiles themselves.

We used two tracks to store and visualize the results of this study on the web: vector polygon data and generalized rasters representing tree crown density. At the native spatial resolution of 50 cm, the map shows the full-resolution tree polygon dataset. At lower-spatial-resolution zoom levels, rasterized representations of tree density are shown. Visualizing generalized rasters in place of vector polygons improves performance substantially. As users zoom in to higher spatial resolutions, the raster layer fades away and is replaced by the full-resolution polygon layer. Once zoomed far enough to resolve individual polygons, users can click to select a polygon to show a map overlay containing various properties of the tree, as well as the date on which the source imagery was acquired and a link to preview the source imagery.

Rainfall data

We used the rainfall data of Funk et al. to estimate annual rainfall at 5.6-m grids[46]. We averaged the available data from 1982 to 2017 and extracted the mean annual rainfall for each mapped tree and bilinearly interpolated it to 100 × 100-m resolution. The rainfall data were also used to classify the study area into mean annual precipitation zones: hyper-arid from 0–150 mm year^{-1}, arid from 150–300 mm year^{-1}, semi-arid from 300–600 mm year^{-1} and sub-humid from 600–1,000 mm year^{-1} zones. The rainfall data are found at https://data.chc.ucsb.edu/products/CHIRPS-2.0/africa_monthly/ (ref. [46]).

Data availability

The viewer can be accessed at https://trees.pgc.umn.edu/app. The Funk et al. rainfall data[46] are freely available at https://data.chc.ucsb.edu/products/CHIRPS-2.0/africa_monthly/. Commercial very-high-resolution satellite images were acquired through NASA under the NextView Imagery End User License Agreement. The copyright remains at Maxar, Inc. and redistribution is not possible. However, the derived products produced by this study are publicly available at the Oak Ridge National Laboratory's Distributed Active Archive Center: https://doi.org/10.3334/ORNLDAAC/2117. Please contact C.T., M.B. or P.H. for more specific requests. A detailed description of our processed data for the 95,402 selected mosaic images, including output data, specific cutlines affected by aerosol optical depth and cloud cover; data distributions for year, month, solar azimuth angle and off-nadir angle for each UTM zone segment in our study, can also be found at https://doi.org/10.3334/ORNLDAAC/2117.

Code availability

The tree detection framework based on U-Net is publicly available at https://zenodo.org/record/3978185. Please contact A.K., C.I., M.B. or J.M. for support and more information.

References

1. Brandt, M. et al. An unexpectedly large count of trees in the West African Sahara and Sahel. *Nature* **587**, 78–82 (2020).

2. Axelsson, C. R. & Hanan, N. P. Patterns in woody vegetation structure across African savannas. *Biogeosciences* **14**, 3239–3252 (2017).

3. Crowther, T. W. et al. Mapping tree density at a global scale. *Nature* **525**, 201–205 (2015).

4. Bastin, J.-F. et al. The extent of forest in dryland biomes. *Science* **356**, 635–638 (2017).

5. Hansen, M. C. et al. High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).

6. Avitabile, V. et al. An integrated pan-tropical biomass map using multiple reference datasets. *Glob. Change Biol.* **22**, 1406–1420 (2016).

7. Baccini, A. et al. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nat. Clim. Change* **2**, 182–185 (2012).

8. Bouvet, A. et al. An above-ground biomass map of African savannahs and woodlands at 25 m resolution derived from ALOS PALSAR. *Remote Sens. Environ.* **206**, 156–173 (2018).

9. Hanan, N. P. et al. Gridded estimates of woody cover and biomass across sub-Saharan Africa, 2000-2004. ORNL DAAC https://doi.org/10.3334/ORNLDAAC/1777 (2020).

10. Saatchi, S. S. et al. Benchmark map of forest carbon stocks in tropical regions across three continents. *Proc. Natl Acad. Sci. USA* **108**, 9899–9904 (2011).

11. Santoro, M. et al. The global forest above-ground biomass pool for 2010 estimated from high-resolution satellite observations. *Earth Syst. Sci. Data* **13**, 3927–3950 (2021).

12. Skole, D. L., Samek, J. H., Dieng, M. & Mbow, C. The contribution of trees outside of forests to landscape carbon and climate change mitigation in West Africa. *Forests* **12**, 1652 (2021).

13. Harris, N. L. et al. Global maps of twenty-first century forest carbon fluxes. *Nat. Clim. Change* **11**, 234–240 (2021).

14. de Foresta, H. et al. *Towards the Assessment of Trees Outside Forests: A Thematic Report Prepared in the Framework of the Global Forest*

Resources Assessment. Forest Resources Assessment Working Paper 183 (FAO, 2013).

15. Bayala, J. et al. Parklands for buffering climate risk and sustaining agricultural production in the Sahel of West Africa. *Curr. Opin. Environ. Sustain.* **6**, 28–34 (2014).

16. Stringer, L. C. et al. Challenges and opportunities in linking carbon sequestration, livelihoods and ecosystem service provision in drylands. *Environ. Sci. Policy* **19-20**, 121–135 (2012).

17. Brito, J. C. et al. Unravelling biodiversity, evolution and threats to conservation in the Sahara-Sahel. *Biol. Rev.* **89**, 215–231 (2014).

18. Bastin, J.-F. et al. The global tree restoration potential. *Science* **365**, 76–79 (2019).

19. Hiernaux, P. et al. Allometric equations to estimate the dry mass of Sahel woody plants mapped with very-high resolution satellite imagery. *For. Ecol. Manag.* **529**, 120653 (2023).

20. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).

21. Reichstein, M. et al. Deep learning and process understanding for data-driven Earth system science. *Nature* **566**, 195–204 (2019).

22. Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015)* (eds Navab, N. et al.) 234–241 (Springer, 2015).

23. Sitch, S. et al. Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences* **12**, 653–679 (2015).

24. Kirkby, M. Desertification and development: some broader contexts. *J. Arid Environ.* **193**, 104575 (2021).

25. Burrell, A. L., Evans, J. P. & De Kauwe, M. G. Anthropogenic climate change has driven over 5 million km^2 of drylands towards desertification. *Nat. Commun.* **11**, 3853 (2020).

26. Darkoh, M. B. K. The nature, causes and consequences of desertification in the drylands of Africa. *Land Degrad. Dev.* **9**, 1–20 (1998).

27. Ribot, J. C. A history of fear: imagining deforestation in the West African dryland forests. *Glob. Ecol. Biogeogr.* **8**, 291–300 (1999).

28. Fairhead, J. & Leach, M. *Reframing Deforestation*: *Global Analyses and Local Realities: Studies in West Africa* (Routledge, 2003).

29. Reed, J. et al. Trees for life: the ecosystem service contribution of trees to food production and livelihoods in the tropics. *For. Policy Econ.* **84**, 62–71 (2017).

30. Descroix, L. et al. Evolution of surface hydrology in the Sahelo-Sudanian strip: an updated review. *Water* **10**, 748 (2018).

31. Wendling, V. et al. Drought-induced regime shift and resilience of a Sahelian ecohydrosystem. *Environ. Res. Lett.* **14**, 105005 (2019).

32. Skole, D. L. et al. Trees outside of forests as natural climate solutions. *Nat. Clim. Change* **11**, 1013–1016 (2021).

33. O'Connor, D. & Ford, J. Increasing the effectiveness of the "Great Green Wall" as an adaptation to the effects of climate change and desertification in the Sahel. *Sustainability* **6**, 7142–7154 (2014).

34. Mirzabaev, A. et al. Economic efficiency and targeting of the African Great Green Wall. *Nat. Sustain.* **5**, 17–25 (2022).

35. Sacande, M. & Berrahmouni, N. Community participation and ecological criteria for selecting species and restoring natural capital with native species in the Sahel. *Restor. Ecol.* **24**, 479–488 (2016).

36. Brandt, M. et al. Changes in rainfall distribution promote woody foliage production in the Sahel. *Commun. Biol.* **2**, 133 (2019).

37. Wigneron, J.-P. et al. SMOS-IC data record of soil moisture and L-VOD: historical development, applications and perspectives. *Remote Sens. Environ.* **254**, 112238 (2021).

38. Schaaf, C. B. et al. First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sens. Environ.* **83**, 135–148 (2002).

39. Tucker, C. J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **8**, 127–150 (1979).

40. Lyapustin, A., Wang, Y., Korkin, S. & Huang, D. MODIS collection 6 MAIAC algorithm. *Atmos. Meas. Tech.* **11**, 5741–5765 (2018).

41. Hiernaux, P. & Le Houerou, H.-N. Les parcours du Sahel. *Secheresse* **17**, 51–71 (2006).

42. White, F. *The Vegetation of Africa. Natural Resources Research, UNESCO Vol. 20* (UNESCO Press, 1983).

43. Schnell, R. *Introduction à la Phytogéographie des Pays Tropicaux. La Flore et la Végétation de l'Afrique Tropicale* (Gauthier-Villars, 1976).

44. McGroddy, M. E., Daufresne, T. & Hedin, L. Scaling of C:N:P stoichiometry in forests worldwide: implications of terrestrial Redfield-type ratios. *Ecology* **85**, 2390–2401 (2004).

45. Bevington, P. & Robinson, D. K. *Data Reduction & Error Analysis for the Physical Sciences* 3rd edn (McGraw-Hill, 2003).

46. Funk, C. et al. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Sci. Data* **2**, 150066 (2015).

47. Djomo, A. N. & Chimi, C. D. Tree allometric equations for estimation of above, below and total biomass in a tropical moist forest: case study

with application to remote sensing. *For. Ecol. Manag.* **391**, 184–193 (2017).

48. Kuyah, S. et al. Allometric equations for estimating biomass in agricultural landscapes: II. Belowground biomass. *Agric. Ecosyst. Environ.* **158**, 225–234 (2012).

49. Dardel, C. et al. Re-greening Sahel: 30 years of remote sensing data and field observations (Mali, Niger). *Remote Sens. Environ.* **140**, 350–364 (2014).

50. Tucker, C. GIMMS global agricultural monitoring system. NASA https://glam1.gsfc.nasa.gov (2022).

Acknowledgements

We thank Maxar, Inc. for providing commercial satellite data through the NextView Imagery End User License Agreement. M.B. and F.R. were supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 947757 TOFDRY) and a DFF Sapere Aude grant (no. 9064–00049B). A.K., R.F. and C.I. acknowledge support by the Villum Foundation through the project Deep Learning and Remote Sensing for Unlocking Global Ecosystem Resource Dynamics (DeReEco, grant no. 34306). C.I. acknowledges support by the Pioneer Centre for Artificial Intelligence, DNRF grant number P1. NASA supported this work through the Commercial SmallSat Data Acquisition (CDSA) Program. We thank NASA's K. Murphy and A. Hall for funding this work; T. Lee and D. Duffy of NASA for supporting our use of Maxar data; B. Kramer, B. Bode and the entire Blue Waters staff who enabled our extensive use of high-performance computing to map 9,947,310,221 trees and convert them into wood, foliage and root carbon.

Author information

Authors and Affiliations

1. Earth Science Division, NASA Goddard Space Flight Center, Greenbelt, MD, USA

Compton Tucker, Jennifer Small, Katherine Melocik, Jesse Meyer, Scott Sinno, Eric Romero, Erin Glennie, Yasmin Fitts, August Morin, Jorge Pinzon, Devin McClain & Benjamin Poulter

2. Science Systems and Applications, Inc., NASA Goddard Space Flight Center, Greenbelt, MD, USA

Martin Brandt, Pierre Hiernaux, Ankit Kariryaa, Jennifer Small, Katherine Melocik, Jesse Meyer, Scott Sinno, Eric Romero, Erin Glennie, Yasmin Fitts, August Morin, Jorge Pinzon & Devin McClain

3. Department of Geosciences and Natural Resource Management, University of Copenhagen, Copenhagen, Denmark

Martin Brandt, Ankit Kariryaa, Kjeld Rasmussen, Florian Reiner & Rasmus Fensholt

4. Pastoralisme Conseil, Caylus, France

Pierre Hiernaux

5. Department of Computer Science, University of Copenhagen, Copenhagen, Denmark

Ankit Kariryaa & Christian Igel

6. Learning and Environmental Sciences, University of Minnesota, Saint Paul, MN, USA

Paul Morin, Claire Porter & Shane Loeffler

7. Géosciences Environnement Toulouse, Observatoire Midi-Pyrénées, UMR 5563 (CNRS/UPS/IRD/CNES), Toulouse, France

Laurent Kergoat

8. Dan Dicko Dankoulodo University of Maradi, Maradi, Niger

Bil-Assanou Issoufou

9. FAO Subregional Office for West Africa, Dakar, Senegal

Patrice Savadogo

10. ISPA, UMR 1391, INRAE Nouvelle-Aquitaine Bordeaux, Villenave d'Ornon, France

Jean-Pierre Wigneron

11. Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, CE Orme des Merisiers, Gif sur Yvette, France

Philippe Ciais

12. Department of Earth & Environment, Boston University, Boston, MA, USA

Robert Kaufmann & Ranga Myneni

13. Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

Sassan Saatchi

Contributions

C.T., M.B., P.H. and A.K. contributed equally to the paper and should be considered as the first authors. C.T. and M.B. coordinated the study and drafted the manuscript, with support by K.R., R.F. and P.H. P.H., P.S. and B.-A.I. collected, assembled and developed the allometric equations. Satellite data were prepared by J.S., S. Sinno, J.M., C.P. and Y.F. J.P. and J.M. directed the evaluation, which was conducted by K.M., E.R., D.M., A.M., A.K., J.S. and C.T. The tree detection was processed on the Blue Waters supercomputer by J.M., J.S. and S. Sinno. Neural network implementation was provided by A.K. and C.I. Ecosystem models were run and analysed by

B.P. and P.C. LVOD data were prepared by J.-P.W. and M.B. Data were analysed by J.P., E.R., E.G., R.F., J.M., D.M., A.M., A.K., K.M., J.S., M.B., F.R., C.T., C.I., R.K., R.M., S. Saatchi and Y.F. Phenology was provided by Y.F., P.H. and L.K. The viewer was developed by P.M. and S.L.

Corresponding authors

Correspondence to Compton Tucker, Martin Brandt or Pierre Hiernaux.

Ethics declarations

Competing interests

The authors declare no competing financial interests.

Peer review

Peer review information

Nature thanks Jules Bayala and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Planet NDVI time series for 308 trees contrasted with the background NDVI from the same area.

a, Note the separation between the tree NDVI values and the background NDVI for the dry-season months of November to March. **b**, False-colour

tree crown NDVI image from WorldView-3 at 30-cm spatial resolution showing green leaf vegetation in red colours. Individual trees are evident, with green leaf tree crowns with associated shadows. Our mapping of trees with machine learning is based on trees defined as having a tree crown of at least 3.0 m^2 with an associated shadow. Areas of low vegetation do not cast shadows sufficiently to be classified as trees. The areas in **a** and **b** are close to 13° 31′ N × 2° 40′ E.

Extended Data Fig. 2 Allometric equations for converting tree crown areas into dry wood mass, dry root mass and dry foliage mass.

a, Allometric equation based on tree crown area to predict wood mass was established from 698 Sahelian and Sudanian woody plants of 27 species. The power model was fitted using log–log regression. The stored carbon is estimated from dry mass by multiplying with a factor of 0.47. Data were collected in the 0–800 mm $year^{-1}$ long-term precipitation zone (see ref. [19]). The plot also shows the cumulative percentage up to 15 m^2 and 50 m^2 crown area (50% and 88.4%, respectively) of the predicted trees in our study (95% of which had crown areas less than 78 m^2). **b**,**c**, Same as in **a** but for foliage and root mass. The foliage and root allometric equations were established from 900 trees of 26 species and 26 trees of five species, respectively.

Extended Data Fig. 3 Comparison between aboveground carbon-density maps and our data derived from 9,947,310,221 trees.

a, Scatter plots showing the spatial agreement at the pixel level with all datasets aggregated to 1 × 1-km grids, with our data plotted on the *x* axes. **b**, We correlated all 1 × 1-km pixels for each rainfall zone, defined by 100 mm $year^{-1}$ steps, between our aboveground carbon density (foliage + wood) with the estimates and current state-of-the-art aboveground carbon-density maps[6,7,8,9,10,11] and our data. See also Fig. 2a.

Extended Data Fig. 4 Temporal changes in carbon density.

a, Scatter plot between the passive microwave L-VOD[36] and our carbon density (wood + foliage) aggregated to 25 × 25-km grids. **b**, The linear relationship seen in **a** was used to convert annual L-VOD to the unit carbon density. L-VOD aboveground woody carbon density as well as TRENDY models[23] values (all vegetation carbon) were averaged over the study area for the 0–1,000 mm year^{-1} rainfall zone for the period 2011–2019. Correlating L-VOD from the dry season, to avoid the complication of herbaceous vegetation, with our carbon-density map aggregated to 25 × 25-km resolution showed a moderately high level of agreement (r = 0.72); however, the strong scattering especially in low-rainfall areas also showed that the uncertainty was high, impeding the use of L-VOD[37] for local applications in arid areas. Nevertheless, the linear relationship was used to convert L-VOD to the unit carbon density to derive temporal dynamics in carbon density, which showed stable woody carbon stocks during 2010–2019 (about 2.0 Mg C year^{-1}) for this region, without clear trend or inter-annual variations, suggesting that none of droughts, deforestation or restoration had a measurable impact on carbon stocks over the past decade. TRENDY models showed a variety of responses but the ensemble showed a similar behaviour as L-VOD, although when herbaceous vegetation and belowground biomass were included, a variety of different magnitudes resulted.

Extended Data Fig. 5 Seasonal comparison between MODIS NDVI from a 0.25° × 0.25° area and field data collections of dry herbaceous biomass production in kg ha^{-1} near Agoufou, Mali within the MODIS area.

a, The MODIS 8-day time-step data showed the NDVI maximum and minimum range from 2000 to 2021 in the grey-coloured portion, with MODIS NDVI time series by years for 2004 to 2010 and the average MODIS NDVI from 2001 to 2021. **b**, Aboveground herbaceous dry mass biweekly in situ measurements for a rangeland field site in Mali[49] centred at 15.4625° N by 1.4886° W. These showed high inter-annual and intra-annual herbaceous dry mass variability, sharp increases in herbaceous dry mass during the wet season and rapid decreases in early dry season at the few-metres scale. The MODIS NDVI data showed similar temporal trends to the

herbaceous dry-mass variations in **b** for 2004–2010 and put these into context of the 2000–2021 MODIS record for a 0.25° × 0.25° area. Trees was roughly 3% of the total vegetation cover in this area and thus herbaceous vegetation dominates. The species composition in **b** was dominated by annual grasses, such as *Aristida mutabilis*, *Cenchrus biflorus* and *Brachiaria xantholeuca*, and annual dicotyledons, such as *Zornia glochidiata* and *Tribulus terrestris*. The data in **b** were from a 1 × 1-km location and were selected to be representative of the area. The MODIS NDVI data in **a** are available for all users, with instructions for use for the MODIS record from 2000 to the present[50]. Our tree crown data provide the means to separate primary production into herbaceous and tree fractions for semi-arid areas and will improve carbon residence understanding in areas of mixed tree and herbaceous vegetation.

Extended Data Fig. 6 Forest fractions of different TRENDY23 models.

The figure shows the percentage of forest areas assumed by each model along the rainfall gradient. Woody biomass in most models mainly comes from predefined forest areas and consequently results in a high degree of variation in the semi-arid area of our study. This is an example of the utility of our tree-mapping results for more accurate depiction of semi-arid trees for numerical simulation modelling.

Extended Data Fig. 7 Candidate and selected satellite images for our analysis.

a, 326,523 Maxar images covering 90,097,073 km^2 from November to March were available for our study and were acquired from 2002 to 2020. **b**, The distribution by month of all satellite data in **a**. The selected imagery for processing (see Methods) by year totalled 94,502 images that covered 9,685,324 km^2 with 87% of the satellite data from 2010 to 2020 (**c**) and 94% of the selected images were from the early dry-season months of November, December and January (**d**). We had a 9.3:1 area ratio of available imagery to selected input data for analysis. See also Extended Data Table 1a.

Extended Data Fig. 8 Evaluation of the tree crown mapping omission and commission errors.

The performance of the tree crown predictions was evaluated for 1,028 randomly selected 512 × 256-pixel areas from UTM zones 28–37, with approximately 100 areas for each UTM zone. The patches were extracted with an 80% bias towards precipitation[46] above 200 mm $year^{-1}$, as most tree identifications are above the 200 mm $year^{-1}$ isoline. 178 areas were excluded from evaluation because the rainfall was >1,000 mm $year^{-1}$ or the areas were devoid of trees. A total of 50,740 trees were evaluated. **a**, Class breakdown of our 3,765 omission and commission errors. **b**, Class numbers of all the trees evaluated. **c**, Summarizes the number of trees, errors of commission and omission and the number of trees by error classes. The highest percentage of binned-error classes, from 15% to >25%, occurred in areas with few trees and resulted from mixed tree and bush confusion for <8% of all trees evaluated. By contrast, the lowest binned-error classes, from 0% to 15%, had 92% of all trees evaluated. In the 850 patches, the number of trees ranged from one tree to 326 trees, with a total of 50,740 trees evaluated and 3,765 errors identified. Overall, the commission and omission error rates were 4.9% and 2.7%, respectively, a net uncertainty of 2.2%.

Extended Data Fig. 9 Tree crown and carbon errors.

The RMSEs of our predicted crown areas calculated on external validation data from ref. [1], binned on the basis of the 50th quantiles of the hand-labelled crown areas. In 78 plots of 256 × 256 pixels, the hand-labelled set contained 5,925 trees and the system delineated 5,915 trees. These crown areas were matched using inner spatial join. Multiple overlapping hand-labelled or predicted tree crown areas were merged into multi-polygons before calculating the RMSE. The crown areas of missed tress counted as errors. For calculating the corresponding RMSE of predicted carbon, we relied on the allometric equations given in Extended Data Fig. 2a–c. The abscissa has a logarithmic scale and 95% of our 9.9 billion tree crowns had crown areas <78 m^2 (Fig. 3b).

Extended Data Fig. 10 Flow chart of the different components of our paper.

We show the satellite data used, how these data were organized for processing with our machine learning software and the requisite training data, and how the resulting segmentation of ten-billion-tree crown area resulted. We then show the conversion of tree crown area in tree wood, root and leaf carbon at the tree level for ten billion trees. Last, we show the use of our viewer to enable use of the data we produced, from tree (1) to tree (10^{10}).

Extended Data Table 1 The November to March Maxar satellite images considered and the number selected for our analysis by UTM zone

Supplementary information

Supplementary Information

This file contains Supplementary Fig. 1a–j and Supplementary Data figure information

Peer Review File

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from

the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Tucker, C., Brandt, M., Hiernaux, P. *et al.* Sub-continental-scale carbon stocks of individual trees in African drylands. *Nature* **615**, 80–86 (2023). https://doi.org/10.1038/s41586-022-05653-6

- Received: 16 December 2021
- Accepted: 13 December 2022
- Published: 01 March 2023
- Issue Date: 02 March 2023
- DOI: https://doi.org/10.1038/s41586-022-05653-6

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

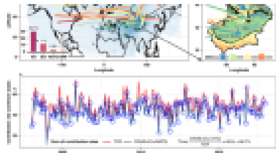
Provided by the Springer Nature SharedIt content-sharing initiative

Carbon stocks of billions of individual African dryland trees estimated

- Jules Bayala
- Meine van Noordwijk

Nature News & Views 01 Mar 2023

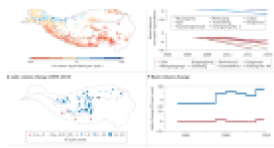
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-022-05653-6



Drying in the low-latitude Atlantic Ocean contributed to terrestrial water storage depletion across Eurasia

06 April 2022

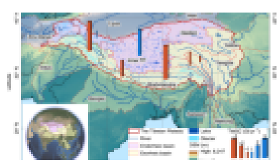
Zexi Shen, Qiang Zhang, … Wenhuan Wu



The imbalance of the Asian water tower

07 June 2022

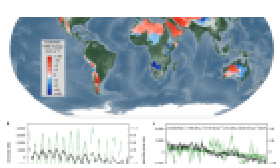
Tandong Yao, Tobias Bolch, … Ping Zhao



Climate change threatens terrestrial water storage over the Tibetan Plateau

15 August 2022

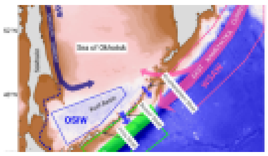
Xueying Li, Di Long, … Guangqian Wang



Recent global decline in endorheic basin water storages

30 November 2018

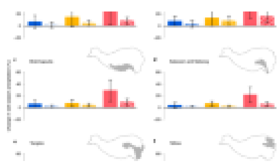
Jida Wang, Chunqiao Song, … Yoshihide Wada



Weakened overturning and tide control the properties of Oyashio Intermediate Water, a key water mass in the North Pacific

15 July 2021

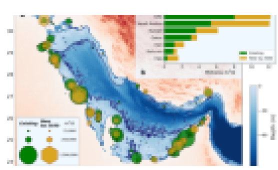
Vigan Mensah & Kay. I. Ohshima



Atmospheric dynamic constraints on Tibetan Plateau freshwater under Paris climate targets

11 January 2021

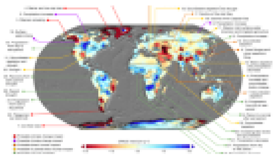
Tao Wang, Yutong Zhao, … Tandong Yao



Long-term, basin-scale salinity impacts from desalination in the Arabian/Persian Gulf

29 November 2022

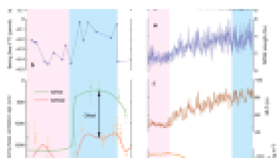
Francesco Paparella, Daniele D’Agostino & John A. Burt



[Emerging trends in global freshwater availability]

16 May 2018

M. Rodell, J. S. Famiglietti, … M.-H. Lo



[Enhanced North Pacific deep-ocean stratification by stronger intermediate water formation during Heinrich Stadial 1]

08 February 2019

X. Gong, L. Lembke-Jene, … X. F. Shi

- Article
- Open Access
- Published: 01 March 2023

Oceanic climate changes threaten the sustainability of Asia’s water tower

- Qiang Zhang ,

- Zexi Shen ,
- Yadu Pokhrel,
- Daniel Farinotti,
- Vijay P. Singh,
- Chong-Yu Xu,
- Wenhuan Wu &
- …
- Gang Wang

Show authors

Nature **volume 615**, pages 87–93 (2023)

- 4382 Accesses
- 66 Altmetric
- Metrics details

Abstract

Water resources sustainability in High Mountain Asia (HMA) surrounding the Tibetan Plateau (TP)—known as Asia's water tower—has triggered widespread concerns because HMA protects millions of people against water stress[1,2]. However, the mechanisms behind the heterogeneous trends observed in terrestrial water storage (TWS) over the TP remain poorly understood. Here we use a Lagrangian particle dispersion model and satellite observations to attribute about 1 Gt of monthly TWS decline in the southern TP during 2003–2016 to westerlies-carried deficit in precipitation minus evaporation (PME) from the southeast North Atlantic. We further show that HMA blocks the propagation of PME deficit into the central TP, causing a monthly TWS increase by about 0.5 Gt. Furthermore, warming-induced snow and glacial melt as well as drying-induced TWS depletion in HMA weaken the blocking of HMA's mountains, causing persistent northward expansion of the TP's TWS deficit since 2009. Future projections under two emissions scenarios verified by satellite observations during 2020–2021 indicate that, by the end of the twenty-first century, up to 84% (for scenario

SSP245) and 97% (for scenario SSP585) of the TP could be afflicted by TWS deficits. Our findings indicate a trajectory towards unsustainable water systems in HMA that could exacerbate downstream water stress.

Main

As a vital source of meltwater from seasonal snowpack and thousands of glaciers, HMA surrounding the TP, also known as Asia's water tower, protects approximately 800 million people against water stress[1,2]. Rapidly growing Asian economies that house the largest populations of the world, particularly China, India and Pakistan, rely heavily on the meltwater from HMA and are hence particularly vulnerable to increasing water stress owing to changing freshwater availability across the region[1]. Past decades witnessed exacerbated water stress in these countries by virtue of declining regional TWS[3,4,5,6,7] and rising water demands[1,8,9]. These supply–demand imbalances have been increasingly undermining regional food security and sociopolitical stability, which are expected to deteriorate under a warming climate[10,11,12,13,14]. As a drought-resilient source of water, meltwater from glaciers help mitigate water stress for downstream areas[1]. However, this water source is not sustainable and accelerating glacier mass loss may lead to a cascade of unintended and detrimental environmental and ecological consequences[15,16,17,18,19,20].

Under a warming climate, water availability in HMA is increasingly being affected by accelerating ice mass loss[3,8,19,20] and warming-induced variations in water availability exhibit high spatial variability[8]. During 2000–2019, HMA's glacier loss accounted for about 19% of the global glacier mass loss of 267 ± 16 Gt per year (ref. [19]). This resulted in a decline in TWS in HMA, reducing water availability and exacerbating water stress in the Indus and Brahmaputra river basins[2,5,8]. Meanwhile, concurrent droughts have been observed in the Aral, Chu–Issyk-Kul and Balkhash basins[1] owing to PME deficit from the North Atlantic[7]. By contrast, increased precipitation in the Sanjiangyuan—the headwaters of the Yellow, Yangtze and Lancang rivers—has caused increased TWS over the central TP[5,8].

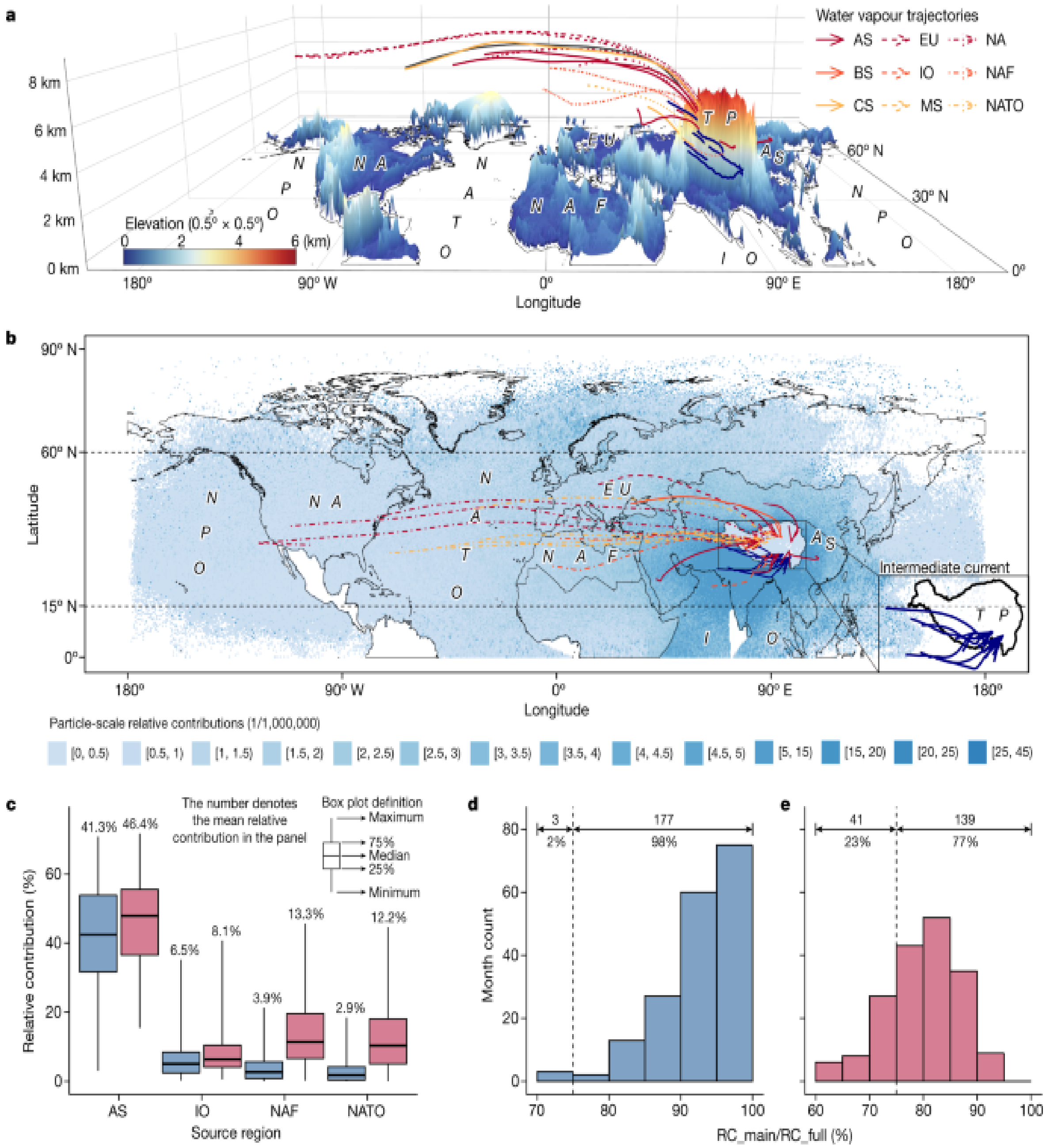
The unprecedented changes in HMA's water systems have been reported in numerous studies[3,8,19,20]. However, the atmospheric mechanisms causing the distinct and regionally varying TWS changes are not well understood[5]. Westerlies and the Indian monsoon bring water vapour recharge for precipitation[21,22,23], and decreasing precipitation in the Himalayas and increasing precipitation in the Pamirs observed during 1979–2010 were attributed to a weakened Indian monsoon[24] and enhanced westerlies[25,26], respectively. Atmospheric circulations also transmit meteorological droughts from the source regions of water vapour[7,27,28], and mid-latitude westerlies have also been identified as potential drivers for HMA's glacier shrinkage[29]. At the same time, the PME deficit transmitted by the westerlies caused TWS deficits across mid-latitude Eurasia but not in the central TP[7]. Given that only limited water vapour reaches the central TP[26], we hypothesize that the TWS increase over the central TP can be attributed to a blocking of the propagation of PME deficits by HMA's high elevations. Because glacier melting dominated TWS changes in HMA[19,30], we argue that shrinking glaciers and changes in snow cover might have modulated the thermodynamic conditions in HMA and resulted in a damping of HMA's blocking.

Analysis of water vapour into the TP

By backtracking water vapour using the FLEXPART model (see Methods), we identify the North Atlantic and the Indian Ocean as the main oceanic water vapour sources for the TP (Fig. 1). When accounting for on-route changes (see Methods), we find long-term average relative contributions of about 42% from Asia, about 7% from the Indian Ocean, about 4% from North Africa and about 3% from the North Atlantic (Fig. 1c). Disregarding on-route changes does not markedly alter the outcomes (Fig. 1c). Crucially, during 98% of the study period (77% when disregarding on-route changes), the above four source regions account for more than 75% of the total water vapour input to the TP (Fig. 1d,e). Distance to the TP has a negative influence on the contributions from source regions to the TP, even though the ocean is the primary continental water source region from a large-scale circulation standpoint. The oceans (for example, the Indian Ocean and the North Atlantic) recharge the water vapour of air particles over the lands (for

example, Asia and Africa) that further transit water vapour into the TP (Supplementary Fig. 1) by means of two main trajectories: the first is from the North Atlantic to the western TP and is mediated by the westerlies (Fig. 1a,b), whereas the second is an intermediate current (see Fig. 1a,b) that brings water vapour from the North Atlantic and the Indian Ocean into the southern TP. Relative contributions from Asia are generally the highest owing to the geographical proximity to the TP. When on-route changes are disregarded, relative contributions from the North Atlantic are larger than those from the Indian Ocean (Fig. 1c), raising the question as to whether the TWS in the southern TP mountains is primarily influenced by variations in water vapour input from the Indian Ocean or the North Atlantic.

Fig. 1: Analysis of water vapour into the TP.



a, The clustered trajectories from source regions to the TP during 2003–2017. **b**, Particle-scale water vapour relative contributions from all source regions to the TP. Intermediate current denotes the current connecting westerlies and the Indian monsoon. **c**, Relative contributions of water vapour from the first to the fourth source regions. **d**,**e** Combined relative contributions (RC_main) for Asia, India Ocean, North Africa and North Atlantic as compared with the total relative contribution of all considered source regions (RC_full). The month count denotes the number of months

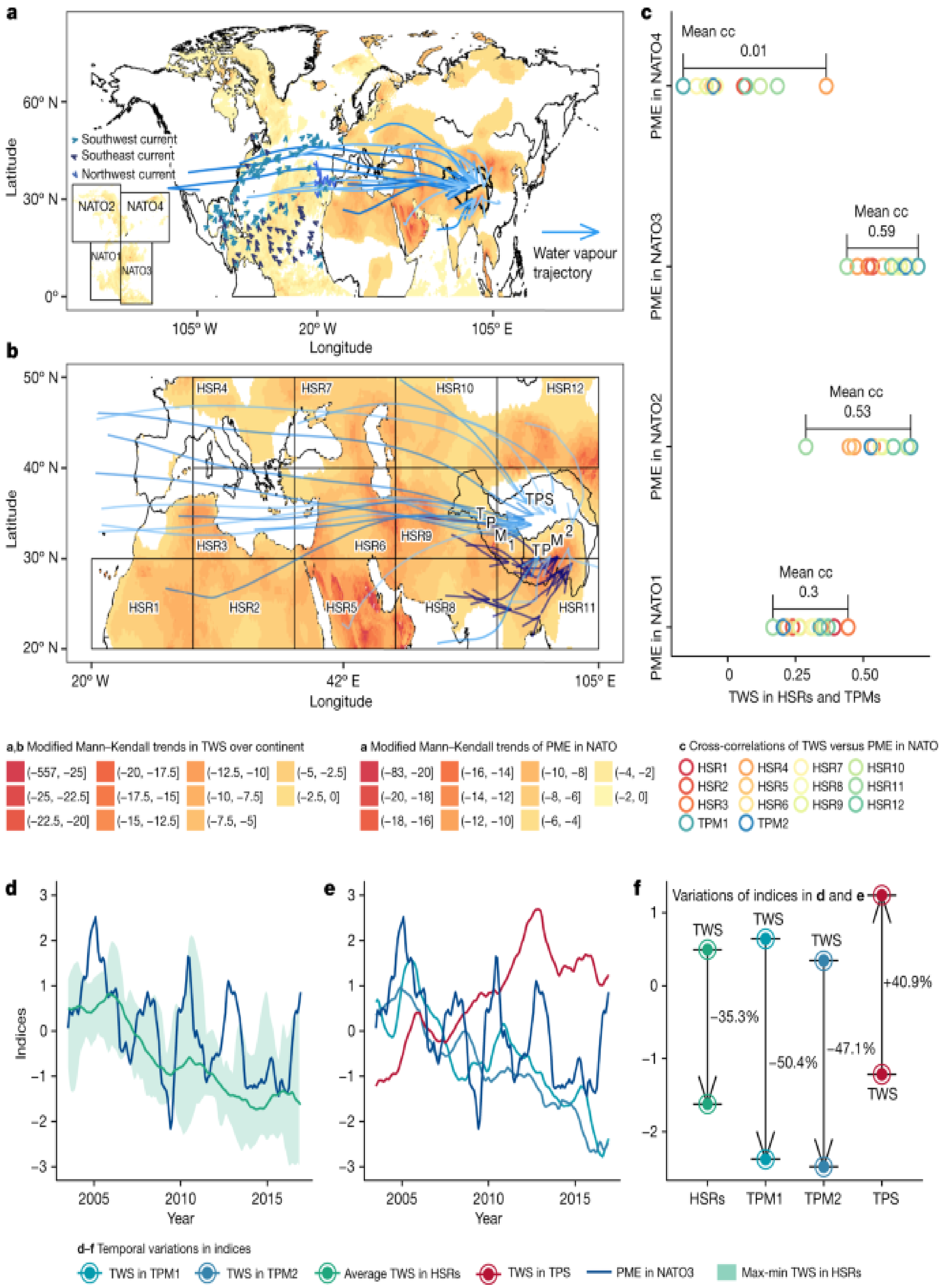
when RC_main/RC_full is in each interval. For panels **c**–**e**, blue colour denotes accounting for on-route changes and pink colour denotes disregarding on-route changes, in which ‘on-route’ means the area between source regions and the TP. The continental map data hereafter are based on country-scale world map data[34]. The TP map data[35] hereafter are acquired from the National Tibetan Plateau Data Center, China. The figure is plotted using R (ref. [36]). The results used to generate the figure are available through Zenodo[37]. AS, Asia (except TP); BS, Black Sea; CS, Caspian Sea; EU, Europe (including Russia); IO, Indian Ocean; MS, Mediterranean Sea; NA, North America; NAF, North Africa; NATO, North Atlantic Ocean; NPO, North Pacific Ocean.

Our results indicate that the TWS changes along the above trajectories are influenced more prominently by the PME deficits in the North Atlantic than by PME variations in the Indian Ocean and local PME variations. On the basis of monthly relative contributions, we find the water vapour input from the Indian Ocean to be dominant during June–October, whereas that from the North Atlantic dominates mainly during November–May (Supplementary Figs. 2 and 3). Seasonal shifts within the source regions complicate the PME pattern along the routes. Poor correlations are detected between (1) PME across three large-scale subregions (SR1–SR3; see Supplementary Fig. 4a–c) along the trajectories and (2) PME in regions of the North Atlantic and the Indian Ocean (Supplementary Fig. 4d–f). The TWS changes can in turn be attributed to the PME deficit transmitted from the water vapour source regions to the TP[7]. The TWS in SR1–SR3, for example, is strongly related to the PME deficit over the northwest and southeast North Atlantic (correlation coefficients of 0.47–0.60, $P < 0.01$; Supplementary Fig. 4g–i). By contrast, TWS variations in SR1–SR3 are poorly correlated with PME over the Indian Ocean and local PME variations (correlation coefficients between −0.07 and 0.40; Supplementary Figs. 4g–i and 5). Furthermore, relationships between TWS over SR1–SR3 and PME over the Indian Ocean are becoming weaker with decreasing distance to the TP (Supplementary Fig. 4g–i). Instead, PME over the northwest and southeast North Atlantic have persistently notable correlations with TWS over SR1–SR3 (Supplementary Fig. 4g–i). Thus, for further analysis, we focus primarily on the impact on TWS over the TP induced by PME deficit over the North Atlantic.

Anomalous TWS increase in the central TP

Covariance analysis suggests a distinct water vapour route that transmits PME deficits from the southeast North Atlantic to the TP. This route is characterized by a synchronous decrease in TWS and spatial overlap with eastward water vapour trajectories (Extended Data Fig. 1). The PME deficit from the low-latitude North Atlantic causes a synchronous TWS deficit over mid-latitude (30 °N–60° N) Eurasia[7]. Focusing on regions within 20° N–50° N and considering optimum lags (see Methods), we show high cross-correlations between the PME deficit over the southeast North Atlantic and the TWS changes in 12 out of 14 subregions discerned along the water vapour propagation routes (Fig. 2a–d and Supplementary Fig. 6). This finding is further corroborated by the leading mode of the maximum covariance analysis between (1) the PME deficit over the North Atlantic and the Indian Ocean and (2) TWS across Eurasia, with an explained covariance ratio of 27% (Extended Data Fig. 1a).

Fig. 2: Anomalous TWS increase over the central TP.


a
60° N
30° N
0°
Latitude
105° W
20° W
105° E
Longitude
Southwest current
Southeast current
Northwest current
NATO2
NATO4
NATO1
NATO3
Water vapour trajectory
b
50° N
40° N
30° N
20° N
Latitude
20° W
42° E
105° E
Longitude
HSR1
HSR2
HSR3
HSR4
HSR5
HSR6
HSR7
HSR8
HSR9
HSR10
HSR11
HSR12
TPS
TPM1
TPM2
c
PME in NATO4
PME in NATO3
PME in NATO2
PME in NATO1
Mean cc 0.01
Mean cc 0.59
Mean cc 0.53
Mean cc 0.3
0
0.25
0.50
TWS in HSRs and TPMs
a,b Modified Mann–Kendall trends in TWS over continent
(−557, −25]
(−25, −22.5]
(−22.5, −20]
(−20, −17.5]
(−17.5, −15]
(−15, −12.5]
(−12.5, −10]
(−10, −7.5]
(−7.5, −5]
(−5, −2.5]
(−2.5, 0]
a Modified Mann–Kendall trends of PME in NATO
(−83, −20]
(−20, −18]
(−18, −16]
(−16, −14]
(−14, −12]
(−12, −10]
(−10, −8]
(−8, −6]
(−6, −4]
(−4, −2]
(−2, 0]
c Cross-correlations of TWS versus PME in NATO
HSR1
HSR2
HSR3
TPM1
HSR4
HSR5
HSR6
TPM2
HSR7
HSR8
HSR9
HSR10
HSR11
HSR12
d
Indices
3
2
1
0
−1
−2
−3
2005
2010
2015
Year
e
3
2
1
0
−1
−2
−3
2005
2010
2015
Year
f
Variations of indices in d and e
1
0
−1
−2
TWS
−35.3%
TWS
−50.4%
TWS
−47.1%
+40.9%
TWS
HSRs
TPM1
TPM2
TPS
d–f Temporal variations in indices
TWS in TPM1
TWS in TPM2
Average TWS in HSRs
TWS in TPS
PME in NATO3
Max-min TWS in HSRs

a,**b**, Spatial pattern of the modified Mann–Kendall trends in TWS across Eurasia and PME over oceans. HSR denotes subregions in which TWS has high cross-correlations with PME over NATO3. **c**, Cross-correlations between PME over the North Atlantic and TWS over HSRs across Eurasia and the southern TP mountains (TPM1 and TPM2). Mean cc denotes the average cross-correlation. **d**, Temporal variations in PME over the southeast North Atlantic (NATO3) and average TWS across the HSRs. The coloured band represents the range from maximum to minimum TWS across the HSRs. **e**, Temporal variations in PME over NATO3 and TWS in the TPMs and the central TP. **f**, Variation in TWS amplitude during 2003–2016. Hereafter, TPM1 and TPM2 refer to southwest and southeast TP mountains, respectively, whereas TPS refers to the central TP surface. The continental map data[34] and the map data of the TP[35] in panels **a** and **b** are from public data sources and plotted using R (ref. [36]). The results used to generate the figure are available through Zenodo[37].

Drying over the North Atlantic induced a TWS decrease between mid-latitude Eurasia (20° N–50° N) and the southern TP mountains, but not in the central TP. We detect a decrease in monthly TWS by about 1 Gt (approximately 47%) over the southern TP mountains during 2003–2016, which we attribute to the propagation of a PME deficit from the southeast North Atlantic (Fig. 2e,f), but a synchronous increase in monthly TWS by about 0.5 Gt (approximately 41%) in the central TP (Fig. 2e,f). The latter has been attributed to the TWS gain caused by increased precipitation after a prolonged dry period[5]. Our linear attribution analysis suggests that the increased TWS over the central TP might be related to decreased TWS across the southern TP mountains instead of local increased PME (Extended Data Fig. 2). Furthermore, the cross-correlation analysis suggests that the decreased TWS over the southern TP mountains can be primarily attributed to the propagation of a PME deficit from the southeast North Atlantic (Extended Data Fig. 3m,n). We also detect a notable relation between TWS depletion (for example, glacier mass loss) and the reductions in snow-cover area over the southern TP mountains (Supplementary Fig. 7). In contrast to TWS declines, the reduction in snow-cover area is mainly caused by regional warming (Supplementary Fig. 8). Although glacier and snow meltwater from the southern TP mountains could increase TWS in the central TP[31], such meltwater also replenishes TWS in the basins surrounding

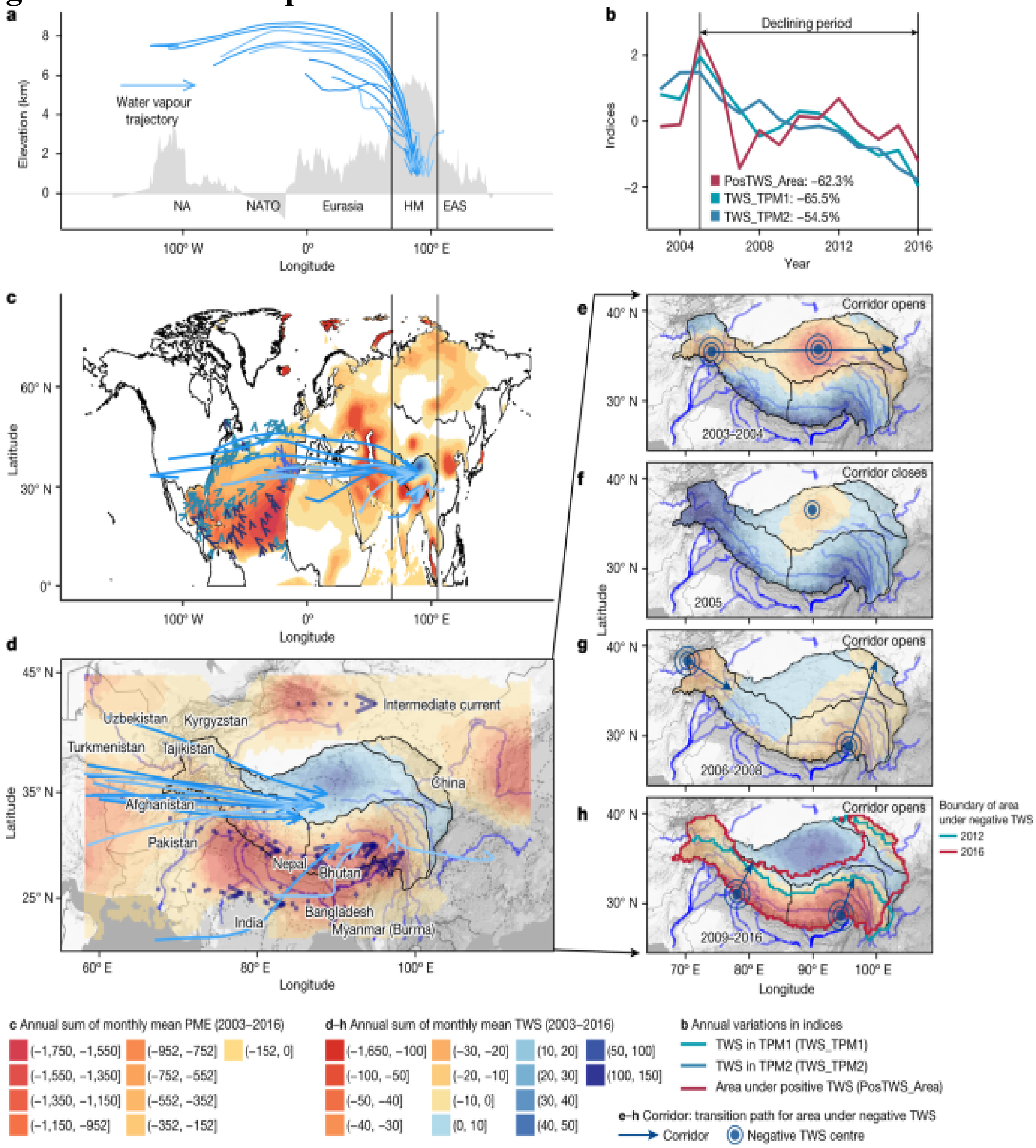
the TP. The endorheic character of the central TP differs from the exorheic nature of the basins situated in the south of the TP (for example, the Indus basin in HSR11; Fig. 2b). This difference could potentially allow for the central TP to accumulate meltwater from the southern TP mountains, thus explaining the observed increase in TWS. However, the fact that the TWS over the largest endorheic basin in China (the Tarim basin in HSR12; Fig. 2b) declines despite the same meltwater supply from the southern TP mountains as for the central TP seems to rule out this hypothesis. Moreover, a comparative analysis for the basins surrounding the central TP (HSR8–HSR12) shows that the surrounding regions are directly exposed to the westerlies and intermediate currents that largely terminate at the southern margin of the central TP (Fig. 2b). This means that the TWS changes in the surrounding basins (Extended Data Fig. 3i–l) are consistent with a propagation of the PME deficit by the westerlies from the southeast North Atlantic, whereas this is not the case for the TWS in the central TP (Extended Data Fig. 3o). Because the southern TP mountains belong to HMA, we suggest that the abnormal increase in TWS in the central TP could be linked to changes in blocking effects caused by HMA's high topography. In particular, this topography dampens the propagation of the PME deficit that emerges in the southeast North Atlantic and we hypothesize that a reduction in the propagation of the PME deficit could have resulted in the observed increase in TWS over the central TP. We investigate this hypothesis further in the next section.

Blocking effects of HMA

FLEXPART simulations suggest that the propagation routes of the PME deficit from the southeast North Atlantic towards the central TP are blocked by HMA's high elevations, causing the routes to split into three trajectories (Figs. 1–3 and Extended Data Fig. 4a). The northern trajectory atmospherically connects the northern TP and southwest TP mountains. The middle trajectory is mainly blocked by the Karakoram and Himalaya mountains, and terminates in the southern parts of the central TP (Fig. 3d and Extended Data Fig. 4a). Finally, the southern trajectory is the intermediate current (Fig. 3d) that runs along the southern side of the Himalaya mountains and moves northward into the TP, along with air

currents from the Indian Ocean. It is blocked by the Himalaya and Nyenchen Tanglha mountains and ends in the southeast TP mountains.

Fig. 3: Northward expansion of TWS deficit in the TP.



a, Elevation profile across the mid-latitude Northern Hemisphere and eastward and northward water vapour trajectories. **b**, Annual sum of the standardized trend items of TWS in the southwest and southeast TP mountains and TWS non-deficit area (PosTWS_Area) in the TP during

2003–2016. **c**,**d** Spatial patterns of annual sums of the monthly mean PME in the North Atlantic and monthly mean TWS across Eurasia during 2003–2016. **e**–**h**, Spatial patterns of the annual sum of TWS during 2003–2004, 2005, 2006–2008 and 2009–2016, respectively. The continental map data[34] and the map data of the TP[35] in panels **c**–**h** are from public data sources and plotted using R (ref. [36]). The background terrain map is from Google Maps acquired using ggmap[38] in R. The results used to generate the figure are available through Zenodo[37]. EAS, East Asia; HM, Himalayan Mountain.

The TWS deficits in the southern TP mountains, together with the warming-induced glacier and snow-cover reduction, modulate the thermodynamic conditions of HMA[32,33]. Particularly, we detect marked teleconnections between (1) air temperatures in the southwest TP mountains and the southern North Atlantic and (2) air temperatures in the southeast TP mountains and the northwest Indian Ocean (Extended Data Fig. 5). These two teleconnections are connected by westerlies and the Indian monsoon, respectively (Fig. 1b). Regional warming in the southern TP has led to a reduction in glacier and snow-cover areas. Studies dedicated to atmospheric flows in mountainous terrain[32,33] indicate that the solar radiation affecting snow-free areas leads to convective thermal internal boundary layers, which in turn give rise to positive buoyancy and thus upslope wind. By contrast, negative buoyancy is promoted over snow-covered and ice-covered areas, inducing katabatic, downslope winds. Here we suggest that the increased occurrence of upslope winds resulting from a reduction of the snow-cover area could have assisted air flow across the HMA main mountain ridges, thus weakening their blocking effect. This weakening of the blocking effect could in turn have resulted in an increased propagation of PME deficits into the northern TP, thus contributing to explaining the observed TWS decline in the northern TP (Supplementary Fig. 9). Such a mechanism would also help explain the correlation identified between TWS changes in the northern TP and the variations in temperature and snow-cover area in the southwest TP mountains (Supplementary Fig. 9b,c).

Furthermore, we detect an annual TWS decrease by about 11 Gt (approximately 66% or roughly 13 $kg\,m^{-3}$) and about 12 Gt (approximately 55% or roughly 12 $kg\,m^{-3}$) in the southwest and southeast TP mountains during 2005–2016, respectively (Fig. 3b). This is partially attributed to

snow-cover reduction (Supplementary Fig. 7). Indeed, the annual average snow-cover areas decline by about 11×10^4 km^2 (approximately 70%) and about 7×10^4 km^2 (approximately 64%) in these two regions, respectively (Extended Data Fig. 6). We suggest that the reductions in snow-cover area and TWS increased upslope winds, thus weakening the blocking effects by assisting air flows (that is, westerlies) to cross HMA. The resulting increase in the propagation of PME deficits into the central TP potentially affects the sustainability of Asia's water tower (Extended Data Fig. 4a). Indeed, the area of the TP affected by a TWS deficit (that is, the area with TWS ≤ 0) increases by approximately 167×10^4 km^2 (about 62%) (Fig. 3b) during 2005–2016.

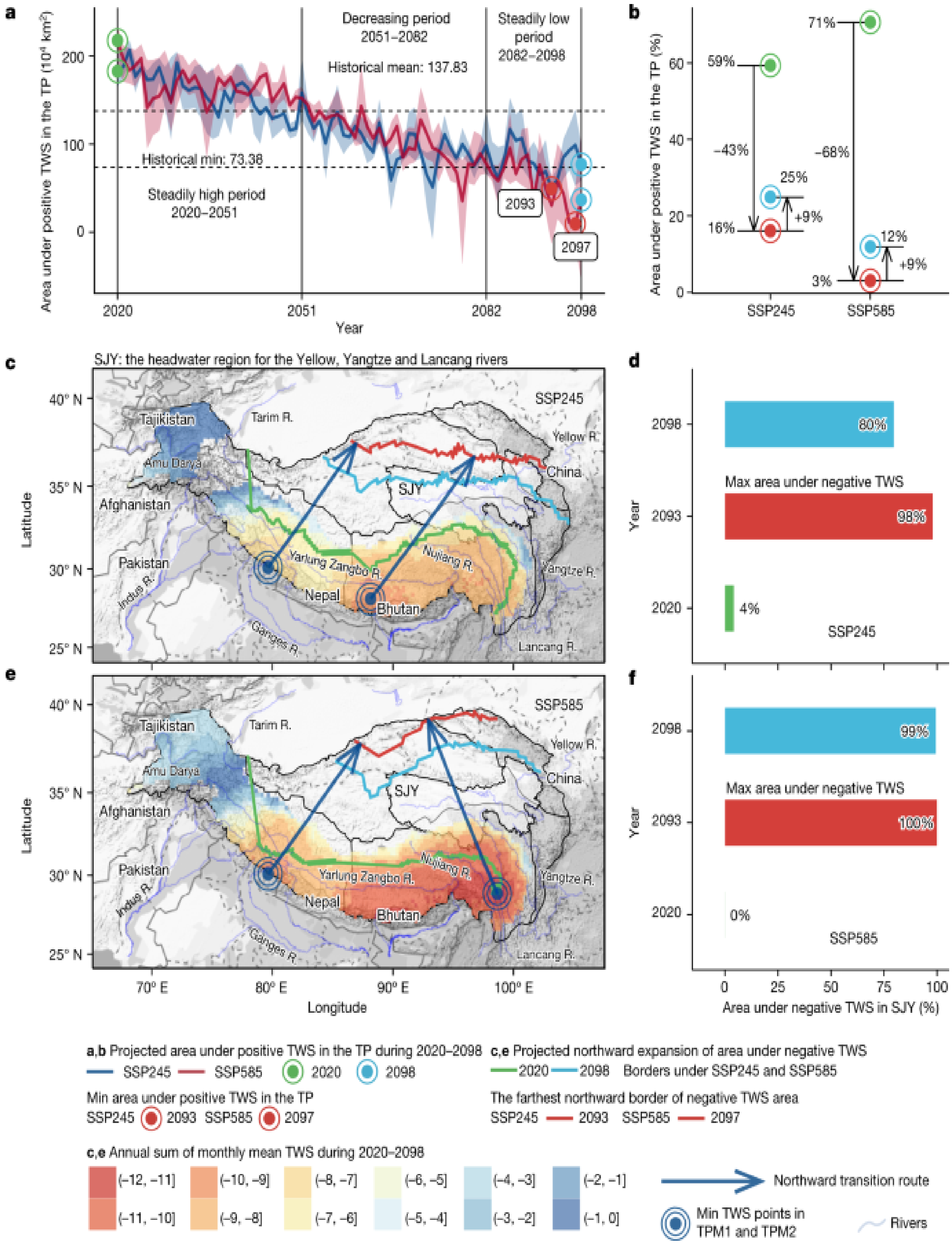
We also find corridors that act as paths for the area with negative TWS to transit from the southern TP mountains to the central TP (Fig. 3e–h). These corridors roughly align with the westerlies or the Indian monsoon (Fig. 3e–h) and open when there is TWS deficit over the southern TP mountains. During 2003–2016, the propagation of the PME deficit along these corridors caused a TWS deficit expansion across the TP of roughly 224×10^4 km^2 (about 73% of the TP) (Fig. 3e–h and Extended Data Fig. 7). In years when the corridor closes (for example, in the year 2005; see Fig. 3f), the related propagation is successfully blocked, preventing a TWS deficit across the central TP. The weakening of the blocking is also visible by the persistent northward expansion of the TWS deficit since 2009—an expansion mediated though middle and southern trajectories (Fig. 3 and Extended Data Figs. 6 and 7). Previous studies have shown that the TWS began to decrease over the central TP in 2013 (refs. [5,30]); however, the cause of such an abrupt TWS decline has not been sufficiently explained. Here we suggest that the persistent northward expansion of the TWS deficit, observed for the TP since 2009, is mainly driven by the westerlies and directly related to the abrupt decrease in 2013 (Figs. 2–3 and Extended Data Figs. 6 and 7).

Future expansion of the TWS deficit

Influenced by a projected warming of air temperatures over oceans by about 1.8–3.9 °C (Supplementary Fig. 10) and by projected climatic drying over the southeast North Atlantic, continued melting of glaciers and snow might contribute to further enhancing upslope winds and thus to further decreasing

the blocking effects of the mountains during the period 2020–2098. Our results indicate that, by the end of the century, this evolution could cause a northward expansion of the area affected by a TWS deficit by up to 258.5 × 10^4 km^2 (84% of the TP) under SSP245 and by up to 298.6 × 10^4 km^2 (97% of the TP) under SSP585 (Fig. 4a, b and Extended Data Fig. 8). Note that the average area of the TP that is not affected by TWS deficits during 2020–2051 is greater than the historical average of 137.8 × 10^4 km^2 (45% of the TP) but then sharply reduces during 2051–2082 (Fig. 4a and Extended Data Fig. 8). The reduction is even more pronounced for the period 2082–2098, in which areas not affected by a TWS deficit are projected to drop to between 27% (SSP245) and 18% (SSP585). This underscores the critical need for reducing carbon emission and for slowing climate warming, otherwise the short-term positive TWS changes in the central TP[8] could reverse. Projected TWS deficits in the southern TP mountains, as well as continuous northward expansion of the TWS deficit over the TP, will probably exacerbate water stress in headwater regions of the main rivers in Asia (Fig. 4c,d). By the end of the century, more than 80% of the Sanjiangyuan will be afflicted by TWS deficit (Fig. 4d,f). These findings imply that the expansion of the area affected by TWS deficit could threaten the sustainability of water supplies from Asia’s water tower, exacerbating water stress in the surroundings and affecting millions of people downstream.

Fig. 4: Futural northward expansion of TWS deficit in the TP.



a, Projected annual variation in the area under positive TWS over the TP in future. For each emissions scenario, the coloured band represents the range from maximum to minimum areas determined by the four methods in this

study. **b**, Amplitude of area under positive TWS changes over the TP in future. **c**,**e**, Projected borders of the areas affected by TWS deficit in 2020, 2093 (2097 under SSP585) and 2098. **d**,**f**, Projected area affected by TWS deficit as a percentage of the total Sanjiangyuan (SJY) area in 2020, 2093 (2097 under SSP585) and 2098. The map data of the TP[35] in panels **c** and **e** are from public data sources and plotted using R (ref. [36]). The background terrain data are from Google Maps acquired using ggmap[38] in R. The results used to generate the figure are available through Zenodo[37].

Methods

Environment input data

All data used in this study are accessible from public databases. For the simulation of water vapour transport (see next section) by the FLEXPART v10.4 model[39,40], ERA-Interim atmospheric files covering 61 vertical levels from 0.1 to 1,000 hPa were downloaded and pre-processed by the flex_extract software (https://www.flexpart.eu/downloads). The spatial and temporal resolutions were 0.75° × 0.75° and 6 h, respectively. Monthly precipitation reanalysis data for 2003–2017 were sourced separately from ECMWF Reanalysis 5th Generation (ERA5)[41] and from the Global Precipitation Climatology Centre (GPCC)[42]. These have a spatial resolution of 0.25° × 0.25°. Reanalysis of monthly evaporation and air temperature for 2003–2017 were sourced from ERA5. Monthly multimodel Coupled Model Intercomparison Project Phase 6 (CMIP6)-based precipitation, evaporation and air temperature data for 2003–2099 were acquired from the Center for Environmental Analysis. The considered CMIP6 models included ACCESS-ESM1.5, BCC-CSM2-MR, CanESM5, GFDL-ESM4, IPSL-CM6A-LR, MIROC6, MRI-ESM2.0 and NorESM2-LM under the r1i1p1f1 format. They were coded as CMIP6 models 1 to 8. The TWS anomaly for 2003–2017 was obtained from the GRACE RL05 Mascon dataset[43] (0.5° × 0.5°), whereas the TWS anomaly during 2020–2021 was taken from the GRACE-FO RL06 Mascon dataset[44] (0.25° × 0.25°). The GRACE RL05 data were interpolated to 0.25° × 0.25° using bilinear interpolation[45]. Monthly mean ocean current data for 1993–2019 were acquired from the Ocean Surface Current Analysis Real-time (OSCAR) dataset[46] and have a spatial resolution

of 0.33° × 0.33°. The snow-cover data (0.05° × 0.05°) were acquired from the National Snow and Ice Data Center[47]. The continental map data used in the study have been generated on the basis of the continental region groups in the country-scale world map data[34]. There is no conflict on the country boundaries because the map shown in the study is continental. The map data for the TP[35] were acquired from the National Tibetan Plateau Data Center, China.

Backward simulation of water vapour source regions for the TP

The Lagrangian particle dispersion model FLEXPART v10.4 simulates forward or backward trajectories of air masses (including their water vapour content) by representing a given air mass as a series of small-scale particles[39]. At the particle scale, the balance between evaporation (e) and precipitation (p) determines the variation in specific humidity (equation (1)). When $e - p > 0$, a particle recharges its water content by evaporation from the surface; when $e - p < 0$, a particle loses its water through precipitation. For a given area, variation in water vapour ($E - P$; uppercase letters denote aggregated quantities) can then be determined as the sum of $e - p$ for all particles in the area (equation (2)),

$$e-p=m\frac{{\rm{d}}q}{{\rm{d}}t} \tag{1}$$

$$E-P=\mathop{\sum }\limits_{i=1}^{n}(e-p) \tag{2}$$

in which $\frac{{\rm{d}}q}{{\rm{d}}t}$ denotes the temporal variation in the specific humidity of a particle, m is the mass of the particle and n is the number of particles over the considered area.

We used FLEXPART v10.4 and ERA-Interim climate reanalysis data to perform backward simulations of water vapour trajectories to the TP. This allows us to determine water vapour source regions. The model was then used to perform monthly simulations of trajectories for up to 500,000 particles. The number of the particles depends on the scale of the target region and computational resources[39,48]. Individual simulations were performed at 6-h time steps for the period 2003–2017, enabling the

determination of water vapour trajectory from source regions to the TP. Given that the residence time of the water vapour in the atmosphere is approximately 10 days (refs. [49,50]), we performed backward simulations from day 1 to day 10 of each month. A *k*-means method[51] was then used to categorize all particle trajectories into 19 out of 20 main water vapour trajectories based on the spatial locations of particles in individual simulations. For the determination of the number of cluster trajectories, see Supplementary Note 1.

Because FLEXPART v10.4 is driven by ERA-Interim and does not adapt to the more recent ERA5 (this choice is due to the fact that ERA5 data supporting the run of FLEXPART v10.4 is not yet accessible to the public user; for details, see https://www.flexpart.eu/flex_extract/Ecmwf/access.html), a comparison was performed between the simulated total water vapour release and ERA5 and GPCC precipitation over the TP to investigate the sensitivity of simulation results[52]. The total water vapour release for particles residing over the TP was calculated by equation (2) when $E - P < 0$ (ref. [53]) and compared with the standardized regional average precipitation obtained from ERA5 and the GPCC. In both cases, comparisons were performed using correlation analysis. It should be noted that the total water vapour release is not equal to precipitation theoretically but it has the positive effect on the formation of the precipitation[39,48,52]. The total released water vapour over the TP is highly correlated to both the GPCC and ERA5 precipitation, with correlation coefficients between 0.85 and 0.86 (P-value < 0.01; see Supplementary Fig. 11). We consider this to be an acceptable accuracy for the FLEXPART results (Supplementary Fig. 11).

On the basis of the initial simulation, 13 primary water vapour source regions were identified: Asia, Indian Ocean, North Africa, North Atlantic, Mediterranean Sea, Europe, Pacific Ocean, Red Sea, Tibet, North America, Caspian Sea, Black Sea and Arctic Ocean. However, because particles can lose or gain water vapour along their track, the water vapour released in the TP can reflect the contribution from both the source regions and the on-route transitions. Thus we calculated the relative contributions in two ways: (1) accounting for or (2) disregarding the on-route changes in water vapour. When accounting for water vapour variations (that is, recharge or loss) along

the transitional routes, the considered air particles show positive variations in water vapour inside the source region (equation (4)), because the negative variations inside the source region indicate that the air particles release water vapour instead of acquiring and transporting water vapour from source regions to the TP. Meanwhile, the considered air particles should also be positive variations in water vapour before reaching the TP when accounting for on-route changes (equation (5)).

$${(E-P)}_{{\rm{route}}}={\sum }_{1}^{r}(e-p)$$
(3)

$${(E-P)}_{{\rm{source}}}={\sum }_{1}^{s}(e-p) > 0$$
(4)

$${(E-P)}_{{\rm{source}}}-{\rm{abs}}\left({\left(E-P\right)}_{{\rm{route}}}\right) > 0$$
(5)

in which $(E-P)_{\rm route}$ and $(E-P)_{\rm source}$ refer to the water vapour variations along the transition routes and over the source regions, respectively. Similarly, r and s refer to the number of particles along the transition routes (r) and over the source regions (s). Because the air particles might gain or lose water vapour, the absolute variations in water vapour along the transition route should be calculated according to equation (5). For particles originating within the TP, $(E-P)_{\rm source}$ should be less than zero, thus recharging water vapour in the TP.

After the selection of particles reaching the TP (equations (3)–(5)), relative water vapour contributions from individual source regions were calculated as follows (equations (6) and (7)).

$${(E-P)}_{i}={\sum }_{1}^{{k}_{i}}(e-p)\left(\,(e-p\right) < 0,\,i=1\,{\rm{to}}\,13$$
(6)

$${{\rm{RC}}}_{i}=\frac{{\left(E-P\right)}_{i}}{{(E-P)}_{{\rm{Tibet}}}}$$
(7)

in which $(E-P)_i$ refers to the total water vapour released over the TP from particles originating from the *i*th source region, $(E-P)_{\mathrm{Tibet}}$ refers to the total water vapour released in the TP and RC_i is the relative water vapour contribution obtained by the TP from the *i*th source region. When accounting for the on-route variation in water vapour, k_i is the number of selected particles (see equations (3)–(5)) from the *i*th source region. When disregarding water vapour on-route variations, k_i refers to the number of all particles from the *i*th source region.

However, there is another method[52] that determines the relative contribution by dividing the total water vapour release over the target region by the moisture gains over the source regions. To avoid impacts induced by the differences between methods on the results, we also calculated the relative contribution using this method. It suggested that both methods come to the same conclusion that Asia, the Indian Ocean, North Africa and the North Atlantic are the four main water vapour source regions for the TP (Supplementary Fig. 12 and Fig. 1c). However, given that the moisture gains over the source regions are far larger than the total water vapour release in the target region, we applied the first method in this study.

Standardized trend item of the index

To avoid the impact of periodicity on correlations between variables, a time series of the variable during January 2003 to June 2017 was first smoothed using a moving average as implemented in the R function decompose. The derived trend series of the variable covers the period July 2003 to December 2016. Then, to ensure that different indices can be compared, the trend series of every variable was standardized over the period 2003–2016 by equation (8).

$$X=\left\{\begin{array}{ll}\frac{{x}_{i}-{\rm{mean}}(x)}{{\rm{sd}}(x)},\quad & x\ne {\rm{TWS}}\\ \frac{{x}_{i}}{{\rm{sd}}(x)},\quad & x={\rm{TWS}}\end{array}\right.$$
(8)

in which *x* denotes the trend series of the variable, *X* denotes the standardized version of index *x*, mean(*x*) and sd(*x*) are the arithmetic average

and standard deviation, respectively, of x over the period 2003–2016 and i refers to the ith month during 2003–2016. Here the variables include PME, TWS, T or Snowcover, in which Snowcover and T denote the snow-covered area of the southern TP mountains and the air temperature, respectively. The relative variation for a standardized index was calculated by dividing the index variation by the maximum amplitude of all standardized indices (see equation (9)). Furthermore, we also provide mass/area change accordingly for every standardized index.

$${{\rm{RV}}}_{m,n}=\frac{{X}_{m}-{X}_{n}}{\max V}\times 100 \%$$
(9)

in which X_m and X_n denote the values at moments m and n in a time series X, $RV_{m,n}$ is the relative variation between X_m and X_n, and $\max V$ denotes the maximum amplitude of all standardized indices in this study. Because the maximum scale for all standardized indices is in the range −3 to 3, we applied $\max V$ as 6 here.

Furthermore, to determine the spatial patterns of the trends, we used the modified Mann–Kendall trend analysis as implemented in the R package modifiedmk. This analysis was conducted at the scale of individual grid cells and covers the period 2003–2016.

Correlation analysis and maximum covariance analysis

Pearson's correlation analysis was used to quantify the relation between standardized indices during 2003–2016. Cross-correlation analysis[54] was applied to evaluate the relation between (1) PME over the North Atlantic and (2) TWS in the 14 small-scale subregions across mid-latitude Eurasia with consideration of the optimum lags. The optimum lag is defined as the lag when the maximum correlations between two indices are detected. Both the Pearson and the cross-correlation analyses were performed using the R package stats.

Maximum covariance analysis was performed to further verify the covariation between spatial matrixes for any pair of indices. The spatial and

temporal coefficients were calculated as follows:

$${\bf{L}}=\left[\begin{array}{ccc}{{\bf{L}}}_{1}(1) & \cdots \, & {{\bf{L}}}_{1}(N)\\ \vdots & \ddots & \vdots \\ {{\bf{L}}}_{{\boldsymbol{m}}}(1) & \cdots \, & {{\bf{L}}}_{m}(N)\end{array}\right]$$
(10)

$${\bf{R}}=\left[\begin{array}{ccc}{{\boldsymbol{R}}}_{1}(1) & \cdots & {{\bf{R}}}_{1}(N)\\ \vdots & \ddots & \vdots \\ {{\bf{R}}}_{{\boldsymbol{q}}}(1) & \cdots & {{\bf{R}}}_{q}(N)\end{array}\right]$$
(11)

$${{\bf{C}}}_{l,r}=\frac{1}{N}{\bf{L}}{{\bf{R}}}^{{\rm{T}}}{\boldsymbol{=}}{\bf{U}}{\boldsymbol{\Sigma }}{{\bf{V}}}^{{\rm{T}}}$$
(12)

$${{\rm{PC}}}_{l,m}={{\bf{U}}}_{m}^{{\rm{T}}}{\bf{L}}$$
(13)

$${{\rm{PC}}}_{r,q}{\boldsymbol{=}}{{\bf{V}}}_{q}^{{\rm{T}}}{\bf{R}}$$
(14)

in which **L** and **R** refer to the spatial matrixes of (1) PME over both the North Atlantic and the Indian Ocean and (2) TWS across Eurasia, respectively, and indices m and q refer to a given cell in the considered field, whereas the index k refers to the time step. $\mathbf{C}_{l,r}$ is the covariance matrix between **L** and **R**. **U** and **V** are the spatial modes, respectively, for **L** and **R**. $\mathrm{PC}_{l,m}$ and $\mathrm{PC}_{r,q}$ are the temporal coefficients for cells m in **L** and q in **R**. The superscript T in equations (10)–(14) denotes the transpose and $\boldsymbol{\Sigma}$ in equation (12) denotes the diagonal matrix. When the maximum correlation analysis is applied to evaluate the relation between T over ocean and land, **L** and **R** also refer to the spatial matrixes of (1) T over the related oceans and (2) T across Eurasia, respectively.

Attribution analysis of TWS variation in the central TP

Previous studies reported that both snowmelt in the southern TP mountains and increased precipitation over the central TP contributed to increased TWS in the central TP[5,31]. However, the balance between precipitation and evaporation was not included in those studies. Here we applied two linear regression models to quantify the contributions of both snowmelt in the southern TP mountains and PME changes in the central TP to TWS changes in the central TP. The models are of the form:

$${\text{TWS}}_{\text{TPS}}={a}_{r}\times {\text{PME}}_{\text{TPS}}+{b}_{r}\times {\text{TWS}}_{r}$$
(15)

in which TWS_r refers to the TWS in a given region r, which is either the southwest TP mountains (TPM1) or the southeast TP mountains (TPM2), TWS_{TPS} and PME_{TPS} denote the TWS and PME in the central TP (TPS), respectively, and a_r and b_r are the regression coefficients determined by linear regression.

Projections of the TWS non-deficit and deficit areas in the TP

The change in area not affected by TWS deficits (hereafter 'TWS non-deficit area') in the TP was determined by identifying grid cells with annual TWS > 0. On the basis of approximately synchronous variations in the TWS non-deficit area and (1) TWS (Fig. 3b) and (2) snow-covered area (Extended Data Fig. 6) in the southwest and southeast TP mountains, a further linear regression model was developed:

$$A={\sum }_{r}({a}_{r}\times {{\rm{TWS}}}_{r}+{b}_{r}\times {{\rm{Snowcover}}}_{r})+\varepsilon $$
(16)

in which A is the TWS non-deficit area in the TP, $Snowcover_r$ is the snow-covered area over region r, a_r and b_r are the regression coefficients and ε is the model residual.

As there are only very few CMIP6 models that simulate TWS_r and $Snowcover_r$ explicitly, three random forest models (equation (17), $m = 1, 2,$

3) were developed for TWS_r and $Snowcover_r$ projections. The models were first trained and validated on the basis of the GRACE-based TWS or the ERA5-based snow-cover area over region r during 2003–2016 and have the form:

$${D}_{r,m}={\rm{RF}}({Y}_{k})$$
(17)

in which Y_k denotes the ERA5-based PME (index $k = 1$) over the southeast North Atlantic, ERA5-based T (index $k = 2$) over the southwest TP mountains or ERA5-based T (index $k = 3$) over the southeast TP mountains during 2003–2016. RF denotes the random forest model. The first random forest model was formed by the dependent variable TWS_r ($D_{r,1}$) and independent variable Y_1 when $m = 1$. The second random forest model was formed by the dependent variable TWS_r ($D_{r,2}$) and independent variables Y_1 and Y_2 in southwest TP mountains or Y_1 and Y_3 in the southeast TP mountains when $m = 2$. The third random forest model was formed by the dependent variable $Snowcover_r$ ($D_{r,3}$) and independent variable Y_2 in the southwest TP mountains or Y_3 in the southeast TP mountains when $m = 3$. The variable TWS_r was projected with ($m = 1$) and without ($m = 2$) considering T, because TWS_r is partially attributed to snow-cover reduction (Supplementary Fig. 7) resulting from local warming (Supplementary Fig. 8).

Equations (16) and (17) were both trained and validated using historical data, reflecting the actual relations between variables (Supplementary Figs. 13–15). These relations were based on the global atmospheric circulation systems (Figs. 1–4) including the westerlies and the Indian monsoon that are generated by the rotation of the Earth and its energy-balance system[55]. Here we project future variations in the TWS non-deficit area in the TP, TWS and snow-covered area in the southern TP mountains using equations (16) and (17) with the CMIP6-modelled data as input.

The distributions of the CMIP6 climatic variables agree well with distributions of the ERA5 climatic variables when $P < 0.05$ (Supplementary Fig. 16). However, given the poor relationships between CMIP6 climate

projections from individual model–scenario combinations and ERA5 climatic variables (Supplementary Figs. 17 and 18), (1) average (equation (18)) and (2) weighting sum (equations (19) and (20)) of climatic indices of CMIP6 models were first determined. These can strongly reduce the uncertainties[56] of CMIP6 projections (Supplementary Figs. 19 and 20) and the average and weighting sum of the climatic indices are obtained as follows:

$${\bar{Y}}_{k,s}=\frac{{\sum }_{i=1}^{8}\left({Y}_{k,i,s}\right)}{8}$$
(18)

$${Y}_{k}=f\left({Y}_{k,i,s},{Z}_{k,i,s}\right)$$
(19)

$${W}_{k,s}={\sum }_{i}({w}_{k,i,s}\times {Y}_{k,i,s}+{c}_{k,i,s}\times {Z}_{k,i,s})+{\varepsilon }_{k,s}$$
(20)

in which s denotes the scenario (either SSP245 or SSP585), ${\bar{Y}}_{k,s}$ denotes the average index for all $Y_{k,i,s}$, $W_{k,s}$ is the weighting sum of $Y_{k,i,s}$ and $Z_{k,i,s}$ denotes either the group of T (index $k = 2$) over the southwest (NATO1) and southeast (NATO3) North Atlantic or the group of T (index $k = 3$) over the four regions in the Indian Ocean (IO1–IO4) derived from the CMIP6 model i under scenario s because air temperature variations over oceans are related to the air temperature over the southern TP mountains (Extended Data Fig. 5). Equation (19) denotes the linear regression model with Y_k (see equation (17)) as the dependent variable and $Y_{k,i,s}$ and $Z_{k,i,s}$ as the independent variables. The weight $w_{k,i,s}$ for $Y_{k,i,s}$, the weight $c_{k,i,s}$ for $Z_{k,i,s}$ and model residuals $\varepsilon_{k,s}$ in equation (20) are determined by equation (19). For the selection of the index i in equation (19), see Supplementary Note 2. Furthermore, the (1) average and (2) weighting sum of TWS_r or $Snowcover_r$ during 2020–2098 under scenario s are projected by equation (17) with ${\bar{Y}}_{k,s}$ (equation (18)) and $W_{k,s}$ (equation (20)) as input, and are defined as (1) ${\bar{D}}_{r,m,s}$ and (2) $WD_{r,m,s}$, respectively.

The annual A during 2020–2098 under scenario s is projected by equation (16) with four combinations of ${\bar{D}}_{r,m,s}$ and $WD_{r,m,s}$

(Supplementary Fig. 21). The average A (see equation (16)) of these projections ($A_{\text{mean},s}$) is determined by equation (21), which has the form:

$${A}_{{\rm{mean}},s}=\frac{{\sum }_{n=1}^{4}{A}_{n,s}}{4}$$
(21)
$${{\rm{Neg}}A}_{{\rm{mean}},s}={A}_{{\rm{TP}}}-{A}_{{\rm{mean}},s}$$
(22)

in which $A_{n,s}$ denotes $A_{1,s}$ projected by equation (16) with $\bar{D}_{r,1,s}$ and $\bar{D}_{r,3,s}$ as inputs, $A_{2,s}$ projected by equation (16) with $\bar{D}_{r,2,s}$ and $\bar{D}_{r,3,s}$ as inputs, $A_{3,s}$ projected by equation (16) with $WD_{r,1,s}$ and $WD_{r,3,s}$ as inputs or $A_{4,s}$ projected by equation (16) with $WD_{r,2,s}$ and $WD_{r,3,s}$ as inputs. Neg$A_{\text{mean},s}$ is the average area affected by TWS deficit (hereafter 'TWS deficit area') during 2020–2098 under scenarios s, computed as the difference between the $A_{\text{mean},s}$ and the total area of the TP (A_{TP}).

Evolution of the area under TWS deficit in the TP

Previous studies suggested that the TWS in the central TP is unpredictable owing to the strong interannual variability[5,31]. By contrast, the TWS in the southwest TP and southeast TP are predictable based on the relation proposed in equation (17). Grid-scale TWS in the southern TP mountains is projected by equation (17) when $m = 1$. The centre points of grid cells, at which the annual sum of the monthly average TWS during 2020–2098 is negative (Fig. 4), on the north margin of the southern TP mountains forms the initial northern border of the TWS deficit area in the TP. Because we only study the spatial variation in the TWS deficit area inside the TP, the southern border of the TP is fixed and the area surrounded by the initial northern border and the southern border of the TP is defined as ini$A$ (Extended Data Fig. 9). The variation in the TWS deficit area (ΔA) was determined by calculating the difference between Neg$A_{\text{mean},y,s}$ and ini$A$:

$${\Delta A}_{y,s}={\rm{ini}}A-{{\rm{Neg}}A}_{{\rm{mean}},y,s}$$
(23)

in which $\text{Neg}A_{\text{mean},y,s}$ is the TWS deficit area in the TP in year y under scenario s. When $\text{ini}A > \text{Neg}A_{\text{mean},y,s}$, the area with TWS < 0 shrinks southward by area $|\Delta A_{y,s}|$. By contrast, when $\text{ini}A < \text{Neg}A_{\text{mean},y,s}$, the area with TWS < 0 expands northward by area $|\Delta A_{y,s}|$.

However, ΔA alone is not enough to determine the areal evolution of the northern border of the TWS < 0 area. Given the stable radial expansion pattern of the TWS < 0 area with two negative centres under the minimum annual TWS sums in the southwest and southeast TP mountains since 2009 (Fig. 3 and Extended Data Fig. 6), future negative centres in these regions are first identified as the cells with the minimum annual sum of the monthly average TWS during 2020–2098 (see Fig. 4c,e) under SSP245 and SSP585 scenarios.

With the aforementioned centres as starting points, we determine the evolving paths as lines between the points forming the initial northern border and the centres in the southern TP mountains (Extended Data Fig. 9). Along the evolving paths, the evolving directions of points forming the initial northern border are thus determined as either northward (for $\text{ini}A < \text{Neg}A_{\text{mean},y,s}$) or southward (for $\text{ini}A > \text{Neg}A_{\text{mean},y,s}$) inside the TP. Furthermore, by setting the expansion steps Δd as 0°, 0.25°, 0.5°, 0.75°, 1°, 2°, 3°, 4°, 5°, 6° and 7°, a fitted line between the pairs of Δd and ΔA is identified in the scatter plot (Extended Data Fig. 9). For a given $\Delta A_{y,s}$, determined by equation (23), a specific Δd could be derived by looking up the scatter plot (Extended Data Fig. 9) using the interpolation method. Once Δd was determined, by moving all of the points forming the initial northern border along the evolving directions, the northern border of the TWS deficit area inside the TP for a specific year and a specific scenario could be determined (Extended Data Fig. 9). We verify this area expansion model by the observed TWS deficit areas during 2009–2016 and find overlapping ratios of 75–91% (Supplementary Fig. 22). Performing the same test for 2020–2021 results in overlapping ratios of 75–86% (Supplementary Fig. 23). We considered these results to be sufficiently accurate for our analysis.

Data availability

All data used in this study were obtained from the public database as described in the 'Environment input data' section. The results used to generate the main figures have been deposited in the public repository Zenodo (https://doi.org/10.5281/zenodo.6790243)[37].

Code availability

All codes used in this study have been deposited in the public repository Zenodo (https://doi.org/10.5281/zenodo.6790243)[37].

References

1. Pritchard, H. D. Asia's shrinking glaciers protect large populations from drought stress. *Nature* **569**, 649–654 (2019).

2. Immerzeel, W. W. et al. Importance and vulnerability of the world's water towers. *Nature* **577**, 364–369 (2020).

3. Kraaijenbrink, P. D. A., Bierkens, M. F. P., Lutz, A. F. & Immerzeel, W. W. Impact of a global temperature rise of 1.5 degrees Celsius on Asia's glaciers. *Nature* **549**, 257–260 (2017).

4. Wang, J. et al. Recent global decline in endorheic basin water storages. *Nat. Geosci.* **11**, 926–932 (2018).

5. Rodell, M. et al. Emerging trends in global freshwater availability. *Nature* **557**, 651–659 (2018).

6. Pokhrel, Y. et al. Global terrestrial water storage and drought severity under climate change. *Nat. Clim. Change* **11**, 226–233 (2021).

7. Shen, Z. et al. Drying in the low-latitude Atlantic Ocean contributed to terrestrial water storage depletion across Eurasia. *Nat. Commun.* **13**, 1849 (2022).

8. Immerzeel, W. W., van Beek, L. P. H. & Bierkens, M. F. P. Climate change will affect the Asian water towers. *Science* **328**, 1382–1385

(2010).

9. Laghari, A. N., Vanham, D. & Rauch, W. The Indus basin in the framework of current and future water resources management. *Hydrol. Earth Syst. Sci.* **16**, 1063–1083 (2012).

10. Schmidhuber, J. & Tubiello, F. N. Global food security under climate change. *Proc. Natl Acad. Sci. USA* **104**, 19703–19708 (2007).

11. Godfray, H. C. J. et al. Food security: the challenge of feeding 9 billion people. *Science* **327**, 812–818 (2010).

12. Kummu, M., Gerten, D., Heinke, J., Konzmann, M. & Varis, O. Climate-driven interannual variability of water scarcity in food production potential: a global analysis. *Hydrol. Earth Syst. Sci.* **18**, 447–461 (2014).

13. Biemans, H. et al. Importance of snow and glacier meltwater for agriculture on the Indo-Gangetic Plain. *Nat Sustain* **2**, 594–601 (2019).

14. IPCC. *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Barros, V. R. et al.) (Cambridge Univ. Press, 2014).

15. Lutz, A. F., Immerzeel, W. W., Shrestha, A. B. & Bierkens, M. F. P. Consistent increase in High Asia's runoff due to increasing glacier melt and precipitation. *Nat. Clim. Change* **4**, 587–592 (2014).

16. Bliss, A., Hock, R. & Radic, V. Global response of glacier runoff to twenty-first century climate change. *J. Geophys. Res. Earth Surf.* **119**, 717–730 (2014).

17. Kääb, A. et al. Massive collapse of two glaciers in western Tibet in 2016 after surge-like instability. *Nature Geosci* **11**, 114–120 (2018).

18. Huss, M. & Hock, R. Global-scale hydrological response to future glacier mass loss. *Nat. Clim. Change* **8**, 135–140 (2018).

19. Hugonnet, R. et al. Accelerated global glacier mass loss in the early twenty-first century. *Nature* **592**, 726–731 (2021).

20. Li, X. et al. Climate change threatens terrestrial water storage over the Tibetan Plateau. *Nat. Clim. Change* **12**, 801–807 (2022).

21. Gimeno, L. et al. Major mechanisms of atmospheric moisture transport and their role in extreme precipitation events. *Annu. Rev. Environ. Resour.* **41**, 117–141 (2016).

22. Gimeno, L. et al. Recent progress on the sources of continental precipitation as revealed by moisture transport analysis. *Earth Sci. Rev.* **201**, 103070 (2020).

23. Shields, C. A. et al. Atmospheric River Tracking Method Intercomparison Project (ARTMIP): project goals and experimental design. *Geosci. Model Dev.* **11**, 2455–2474 (2018).

24. Thompson, L. G. et al. Abrupt tropical climate change: past and present. *Proc. Natl Acad. Sci. USA* **103**, 10536–10543 (2006).

25. Zhao, H. et al. Deuterium excess record in a southern Tibetan ice core and its potential climatic implications. *Clim. Dyn.* **38**, 1791–1803 (2012).

26. Yao, T. et al. Different glacier status with atmospheric circulations in Tibetan Plateau and surroundings. *Nat. Clim. Change* **2**, 663–667 (2012).

27. Herrera-Estrada, J. E. et al. Reduced moisture transport linked to drought propagation across North America. *Geophys. Res. Lett.* **46**, 5243–5253 (2019).

28. Herrera-Estrada, J. E. & Diffenbaugh, N. S. Landfalling droughts: global tracking of moisture deficits from the oceans onto land. *Water Resour. Res.* **56**, e2019WR026877 (2020).

29. Mölg, T., Maussion, F. & Scherer, D. Mid-latitude westerlies as a driver of glacier variability in monsoonal High Asia. *Nat. Clim. Change* **4**,

68–73 (2014).

30. Jacob, T. et al. Recent contributions of glaciers and ice caps to sea level rise. *Nature* **482**, 514–518 (2012).

31. Zhang, G., Xie, H., Kang, S., Yi, D. & Ackley, S. F. Monitoring lake level changes on the Tibetan Plateau using ICESat altimetry data (2003–2009). *Remote Sens. Environ.* **115**, 1733–1742 (2011).

32. Mott, R., Daniels, M. & Lehning, M. Atmospheric flow development and associated changes in turbulent sensible heat flux over a patchy mountain snow cover. *J. Hydrometeorol.* **16**, 1315–1340 (2015).

33. Sauter, T. & Galos, S. P. Effects of local advection on the spatial sensible heat flux variation on a mountain glacier. *Cryosphere* **10**, 2887–2905 (2016).

34. GeoMapApp. World Adm0 map. GeoMapApp http://www.geomapapp.org/GMA/Layers/Shapefiles/Country_Boundaries (2022).

35. Zhang, G. Dataset of river basins map over the TP. National Tibetan Plateau Data Center https://doi.org/10.11888/BaseGeography.tpe.249465.file (2019)

36. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* 2nd edn (Springer, 2016).

37. Zhang, Q. et al. Sharing codes and data for the article entitled "Oceanic climate changes threaten the sustainability of Asia's water tower". Zenodo https://doi.org/10.5281/zenodo.6790243 (2022).

38. Kahle, D., & Wickham, H. ggmap: spatial visualization with ggplot2. *The R Journal* Vol. 5, Issue 1, pp. 144–161. https://journal.r-project.org/archive/2013-1/kahle-wickham.pdf (2013).

39. Pisso, I. et al. The Lagrangian particle dispersion model FLEXPART version 10.4. *Geosci. Model Dev.* **12**, 4955–4997 (2019).

40. Dee, D. P. et al. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**, 553–597 (2011).

41. Hersbach, H. et al. ERA5 monthly averaged data on single levels from 1959 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS) https://doi.org/10.24381/cds.f17050d7 (2019).

42. Schneider, U., Becker, A., Finger, P., Rustemeier, E. & Ziese, M. GPCC full data monthly product version 2020 (at 0.25° at 0.5°, 1.0°, 2.5°): monthly land-surface precipitation from rain-gauges built on GTS-based and historical data. Global Precipitation Climatology Centre https://doi.org/10.5676/DWD_GPCC/FD_M_V2020_025 (2020).

43. Tapley, B. D., Bettadpur, S., Watkins, M. & Reigber, C. The gravity recovery and climate experiment: mission overview and early results. *Geophys. Res. Lett.* **31**, L09607 (2004).

44. Tapley, B. D. et al. Contributions of GRACE to understanding climate change. *Nat. Clim. Change* **9**, 358–369 (2019).

45. Shen, Z. et al. Mining can exacerbate global degradation of dryland. *Geophys. Res. Lett.* **48**, e2021GL094490 (2021).

46. Bonjean, F. & Lagerloef, G. S. E. Diagnostic model and analysis of the surface currents in the tropical Pacific Ocean. *J. Phys. Oceanogr.* **32**, 2938–2954 (2002).

47. Hall, D. K. & Riggs, G. A. MODIS/Terra snow cover monthly L3 global 0.05Deg CMG, version 61 (MOD10CM). NASA National Snow and Ice Data Center Distributed Active Archive Center https://doi.org/10.5067/MODIS/MOD10CM.061 (2021).

48. Ordoñez, P. et al. Climatological moisture sources for the Western North American Monsoon through a Lagrangian approach: their influence on precipitation intensity. *Earth Syst. Dyn.* **10**, 59–72 (2019).

49. Numaguti, A. Origin and recycling processes of precipitating water over the Eurasian continent: experiments using an atmospheric general

circulation model. *J. Geophys. Res.* **104**, 1957–1972 (1999).

50. Gimeno, L. et al. Oceanic and terrestrial sources of continental precipitation. *Rev. Geophys.* **50**, RG4003 (2012).

51. Hartigan, J. A. & Wong, M. A. Algorithm AS 136: a *k*-means clustering algorithm. *J. R. Stat. Soc. Ser. C Appl. Stat.* **28**, 100–108 (1979).

52. Sun, B. & Wang, H. Moisture sources of semiarid grassland in China using the Lagrangian particle model FLEXPART. *J. Clim.* **27**, 2457–2474 (2014).

53. Stohl, A. & James, P. A Lagrangian analysis of the atmospheric branch of the global water cycle. Part I: method description, validation, and demonstration for the August 2002 flooding in central Europe. *J. Hydrometeorol.* **5**, 656–678 (2004).

54. Oppenheim, A. V. & Schafer, R. W. *Discrete-Time Signal Processing* (Prentice Hall, 2009).

55. Wake, B. Earth's energy balance. *Nat. Clim. Change.* **4**, 758 (2014).

56. Shiogama, H. et al. Emergent constraints on future precipitation changes. *Nature* **602**, 612–616 (2022).

Acknowledgements

Q.Z., Z.S., W.W. and G.W. acknowledge the support from the China National Key R&D Program (grant no. 2019YFA0606900), the National Natural Science Foundation of China (grant no. 42041006) and the Engineering Research Center for Water Resources & Ecological Water in Cold and Arid Regions of Xinjiang Uygur Autonomous Region, China (grant no. 2020.A-003). Y.P. acknowledges the support from the National Science Foundation (grant no. 1752729). D.F. acknowledges the support from the Swiss National Science Foundation (grant no. 200021_184634).

Author information

Author notes

1. Qiang Zhang

   Present address: Advanced Interdisciplinary Institute of Environment and Ecology, Beijing Normal University, Zhuhai, China

2. These authors contributed equally: Qiang Zhang, Zexi Shen

Authors and Affiliations

1. State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing, China

   Qiang Zhang, Zexi Shen, Wenhuan Wu & Gang Wang

2. Faculty of Geographical Science, Beijing Normal University, Beijing, China

   Qiang Zhang, Zexi Shen, Wenhuan Wu & Gang Wang

3. Department of Civil and Environmental Engineering, Michigan State University, East Lansing, MI, USA

   Yadu Pokhrel

4. Laboratory of Hydraulics, Hydrology and Glaciology (VAW), ETH Zürich, Zürich, Switzerland

   Daniel Farinotti

5. Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), Birmensdorf, Switzerland

   Daniel Farinotti

6. Department of Biological and Agricultural Engineering, Texas A&M University, College Station, TX, USA

Vijay P. Singh

7. Zachry Department of Civil and Environmental Engineering, Texas A&M University, College Station, TX, USA

Vijay P. Singh

8. National Water and Energy Center, UAE University, Al Ain, United Arab Emirates

Vijay P. Singh

9. Department of Geosciences and Hydrology, University of Oslo, Oslo, Norway

Chong-Yu Xu

Contributions

Q.Z. and Z.S. designed the research and wrote the manuscript. Z.S. performed the analysis. Y.P., D.F., V.P.S., C.-Y.X., W.W. and G.W. discussed and modified the manuscript. Q.Z. and Z.S. contributed equally to the work as the co-first authors and shared equal responsibilities as the corresponding authors.

Corresponding authors

Correspondence to Qiang Zhang or Zexi Shen.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Maximum covariance analysis between oceanic PME and TWS in lands.

a, Denotes time coefficients of PME over ocean and TWS across Eurasia under the first leading mode with the explained variance ratio as 26.68%. 'cor.' quantifies the relation between time coefficients PME over oceans and TWS across Eurasia with the *P*-value of 7.78×10^{-59}. **b**, Depicts spatial coefficients of PME over oceans and TWS across Eurasia under the first leading mode. **c**,**e**, Same as panel **b** but for the North Atlantic (**c**) and the Indian Ocean (**e**). **d**, Presents the modified Mann–Kendall trends in oceanic PME and TWS in lands. The continental world map data[34] in panels **b**–**e** and map data of the TP[35] in panel **d** are acquired from public data sources and plotted using R (ref. [36]). The results used to generate the figure are available through Zenodo[37]. MCA, maximum covariance analysis.

Extended Data Fig. 2 Attribution analysis of the TWS in the central TP.

a,**c**, Temporal variances in TWS and PME in the central TP and TWS in the southwest (TPM1) and southeast (TPM2) TP mountains. **b**,**d**, Include coefficients $a1$, $b1$, $a2$ and $b2$ of the linear regression models in **a** and **c**. The results used to generate the figure are available through Zenodo[37].

Extended Data Fig. 3 PME over NATO3 and continental TWS.

a–**l**, Comparisons between PME over NATO3 and TWS in HSR1–HSR12. **m**,**n**, Same as panels **a**–**l** but for TWS in the southern TP mountains (TPM1 and TPM2). **o**, Same as panels **a**–**l** but for TWS in the TPS. 'ccor' denotes cross-correlation coefficient and the numbers in the parentheses present *P*-values in the cross-correlation analysis. The results used to generate the figure are available through Zenodo[37].

Extended Data Fig. 4 Illustration of the blocking effect by HMA.

a, Mechanism of the interception of the eastward propagation of the PME deficit in the North Atlantic by the southern HMA. **b**, Variation in the interception of the eastward propagation of the PME deficit in the North Atlantic by the southern HMA. The southwest (TPM1) and southeast (TPM2) TP mountains are all in the southern HMA. The terrain texture in panel **a** is acquired from Google and the Tangram Heightmapper (https://tangrams.github.io/heightmapper/) and is plotted using the authorized Adobe Photoshop 2021 plug-in named '3D Map Generator – Terrain'. Panel **b** is plotted using the authorized Adobe Illustrator 2021. This figure is based on the results of this study[37].

Extended Data Fig. 5 Maximum covariance analysis between oceanic and continental *T*.

a, The spatial coefficients of the air temperature (*T*) over oceans and that across Eurasia under the first leading mode. **b**, The time coefficients of the air temperature over oceans and that across the Eurasia under the first leading mode with the explained variance ratio as 27.66%. **c**, Temporal variation in air temperature over the southwest North Atlantic (NATO1) and air temperature in the southwest TP mountains (TPM1). **d**, Temporal variation in air temperature in the northwest Indian Ocean (IO2) and air temperature in the southeast TP mountains (TPM2). The continental world map data[34] and map data of the TP[35] in panel **a** are acquired from public

data sources and plotted using R (ref. [36]). The results used to generate the figure are available through Zenodo[37].

Extended Data Fig. 6 Standardized annual snow-cover area and the area under positive annual TWS in the TP.

The PosTWS_Area denotes the standardized area under positive annual TWS in the TP during 2003–2016. The results used to generate the figure are available through Zenodo[37].

Extended Data Fig. 7 Evolution in the area under negative annual TWS in the TP.

a–**n**, Spatiotemporal transition pattern of the border of the area affected by negative annual sum of TWS in the TP during 2003–2016. **o**, The variation in the area in the TP affected by positive and negative annual sum of TWS. The map data of the TP[35] in panels **a**–**n** are acquired from public data sources and plotted using R (ref. [36]). The results used to generate the figure are available through Zenodo[37].

Extended Data Fig. 8 Future evolution in the area under negative annual TWS in the TP.

a, The average area under annual positive (non-deficit) and negative (deficit) TWS in the TP under SSP245 and SSP585 scenarios, respectively. **b**–**g**, Projected borders for the maximum, mean and minimum areas under negative TWS in the TP under SSP245 scenario (**b**–**d**) and SSP585 scenario (**e**–**g**), respectively. The map data of the TP[35] in panels **b**–**g** are acquired from public data sources and plotted using R (ref. [36]). The background terrain map in panels **b**–**g** is from Google Maps acquired using ggmap[38] in R. The results used to generate the figure are available through Zenodo[37].

Extended Data Fig. 9 Illustration for the projection of borders in the future.

Determination of the projected northward expansion or southward shrinkage area ΔA (**a**) and movement distance Δd (**b**). $\mathrm{Neg}A_{\mathrm{mean},y}$ refers to the averaged projected TWS deficit area by four data inputs in year y (see Methods). For the definition of the initial northern border of the TWS deficit area, see Methods. The map of the TP[35] in panel **b** is acquired from a public data source and plotted using R (ref. [36]). The results used to generate the figure are available through Zenodo[37].

Supplementary information

Supplementary Information

This file contains Supplementary Notes 1 and 2 and Supplementary Figures 1–25.

Peer Review File

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Zhang, Q., Shen, Z., Pokhrel, Y. *et al.* Oceanic climate changes threaten the sustainability of Asia's water tower. *Nature* **615**, 87–93 (2023). https://doi.org/10.1038/s41586-022-05643-8

- Received: 02 July 2022
- Accepted: 09 December 2022
- Published: 01 March 2023
- Issue Date: 02 March 2023
- DOI: https://doi.org/10.1038/s41586-022-05643-8

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

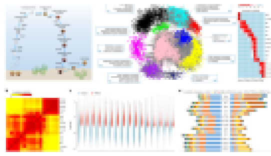
Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

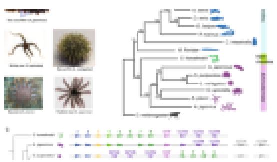
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-022-05643-8



Evolutionary transcriptomics of metazoan biphasic life cycle supports a single intercalation origin of metazoan larvae

16 March 2020

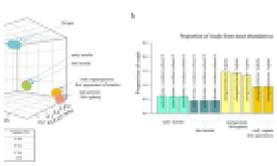
Jing Wang, Lingling Zhang, … Shi Wang



Genomic insights of body plan transitions from bilateral to pentameral symmetry in Echinoderms

10 July 2020

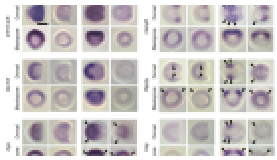
Yongxin Li, Akihito Omori, … Naoki Irie



Unravelling paralogous gene expression dynamics during three-spined stickleback embryogenesis

06 March 2019

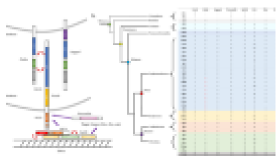
Elisavet Kaitetzidou, Ioanna Katsiadaki, … Elena Sarropoulou



[Genetic regulation of amphioxus somitogenesis informs the evolution of the vertebrate head mesoderm]

01 July 2019

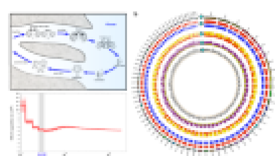
Daniel Aldea, Lucie Subirana, … Hector Escriva



[Comparative and evolutionary analyses reveal conservation and divergence of the notch pathway in lophotrochozoa]

31 May 2021

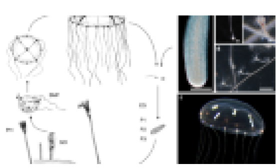
Xin He, Fucun Wu, … Guofan Zhang



[The Chinese mitten crab genome provides insights into adaptive plasticity and developmental regulation]

22 April 2021

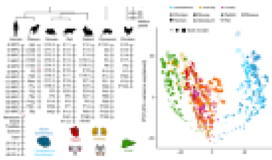
Zhaoxia Cui, Yuan Liu, … Ka Hou Chu



[The genome of the jellyfish Clytia hemisphaerica and the evolution of the cnidarian life-cycle](#)

11 March 2019

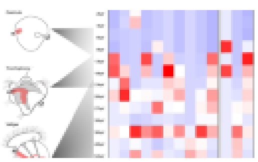
Lucas Leclère, Coralie Horin, … Richard R. Copley



[Gene expression across mammalian organ development](#)

26 June 2019

Margarida Cardoso-Moreira, Jean Halbert, … Henrik Kaessmann



[Non-collinear Hox gene expression in bivalves and the evolution of morphological novelties in mollusks](#)

11 February 2021

David A. Salamanca-Díaz, Andrew D. Calcino, … Andreas Wanninger

- Article
- Open Access
- [Published: 25 January 2023](#)

Annelid functional genomics reveal the origins of bilaterian life cycles

- Francisco M. Martín-Zamora,
- Yan Liang,
- Kero Guynes,
- Allan M. Carrillo-Baltodano,
- Billie E. Davies,
- Rory D. Donnellan,
- Yongkai Tan,
- Giacomo Moggioli,
- Océane Seudre,
- Martin Tran,
- Kate Mortimer,
- Nicholas M. Luscombe,
- Andreas Hejnol,
- Ferdinand Marlétaz &
- …
- José M. Martín-Durán

Show authors

Nature **volume 615**, pages 105–110 (2023)

- 12k Accesses
- 265 Altmetric
- Metrics details

Abstract

Indirect development with an intermediate larva exists in all major animal lineages[1], which makes larvae central to most scenarios of animal evolution[2,3,4,5,6,7,8,9,10,11]. Yet how larvae evolved remains disputed. Here we show that temporal shifts (that is, heterochronies) in trunk formation underpin the diversification of larvae and bilaterian life cycles. We performed chromosome-scale genome sequencing in the annelid *Owenia fusiformis* with transcriptomic and epigenomic profiling during the life cycles of this and two other annelids. We found that trunk development is

deferred to pre-metamorphic stages in the feeding larva of *O. fusiformis* but starts after gastrulation in the non-feeding larva with gradual metamorphosis of *Capitella teleta* and the direct developing embryo of *Dimorphilus gyrociliatus*. Accordingly, the embryos of *O. fusiformis* develop first into an enlarged anterior domain that forms larval tissues and the adult head[12]. Notably, this also occurs in the so-called 'head larvae' of other bilaterians[13,14,15,16,17], with which the *O. fusiformis* larva shows extensive transcriptomic similarities. Together, our findings suggest that the temporal decoupling of head and trunk formation, as maximally observed in head larvae, facilitated larval evolution in Bilateria. This diverges from prevailing scenarios that propose either co-option[9,10] or innovation[11] of gene regulatory programmes to explain larva and adult origins.

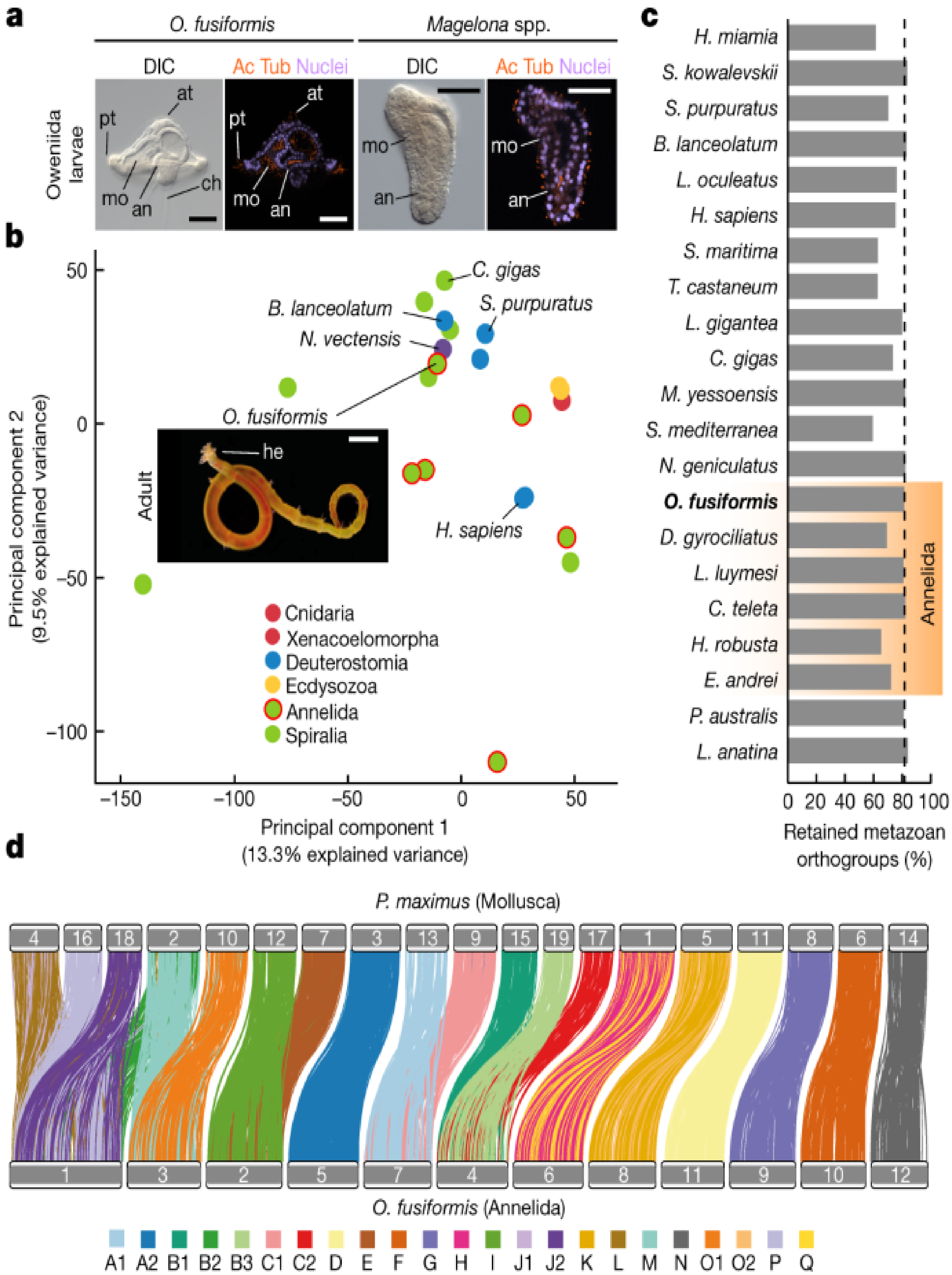
Main

Many animal embryos develop into a larva that metamorphoses into a sexually competent adult[1]. Larvae are morphologically and ecologically diverse, and given their broad phylogenetic distribution, they are central to major scenarios of animal evolution[2,3,4,5,6,7,8,9,10,11]. However, these scenarios dissent on whether larvae are ancestral[2,3,4,5,6] or secondarily evolved[9,10], and on the mechanisms that facilitated the evolution of larvae[2,9,10,11]. Therefore, larval origins—and their importance to explain animal evolution—are still contentious.

The trochophore is a widespread larval type characterized by an apical sensory organ and a pre-oral locomotive ciliary band[18] that is typically assigned to Annelida and Mollusca. Annelids, however, show diverse life cycles and larval morphologies, including species with direct and indirect development and either planktotrophic or lecithotrophic larvae[19]. Notably, the groups Oweniidae and Magelonidae—which form Oweniida, the sister taxon to all other annelids[20]—have distinctive planktotrophic larvae (Fig. 1a and Extended Data Fig. 1a). In particular, the larva of Oweniidae, referred to as 'mitraria'[12], has an enlarged pre-oral region and a bundle of posterior chaetae, as well as a pair of nephridia and a long monociliated ciliary band similar to those of phylogenetically distant larvae of echinoderms and hemichordates[21,22]. Yet oweniids show many developmental characteristics

that are considered ancestral to Annelida and even Spiralia as a whole[23,24], including similarities in larval molecular patterns with other trochophore and bilaterian larvae[22,23,25,26]. Therefore, the diversity of life cycles and larval forms but generally conserved early embryogenesis and adult body plans of Annelida is an excellent model to investigate how larval traits evolve. It is also an ideal model to formulate and assess hypotheses on the origin of larvae and animal life cycles.

Fig. 1: *O. fusiformis* has a distinct larva and a conservatively evolving genome.



a, The larvae of oweniids and magelonids are unlike other annelid larvae. Differential interface contrast (DIC) images and *z*-stack confocal laser

scanning views of a *O. fusiformis* mitraria and a *Magelona* spp. larva stained for nuclei using DAPI and acetylated α-tubulin (Ac Tub). **b**, Principal component analysis of metazoan gene complements demonstrates that *O. fusiformis* clusters with other lineages with conservatively evolving gene complements. See Extended Data Fig. 1e for a fully labelled graph. Inset, image of an adult *O. fusiformis*. **c**, Percentage of retained pre-metazoan and metazoan orthogroups per species. Dotted vertical line represents the value for *O. fusiformis*. A list of species names are provided in Supplementary Table 2. **d**, Karyotypic correspondence between *O. fusiformis* and *Pecten maximus*, which exemplifies the ancestral spiralian chromosome complement. Each colour represents an ancestral bilaterian linkage group. Schematic drawings are not to scale. an, anus; at, apical tuft; ch, chaetae; he, head; mo, mouth; pt, prototroch; tt: telotroch. Scale bars, 50 µm (**a**) or 2.5 mm (**b**).

O. fusiformis has a conserved genome

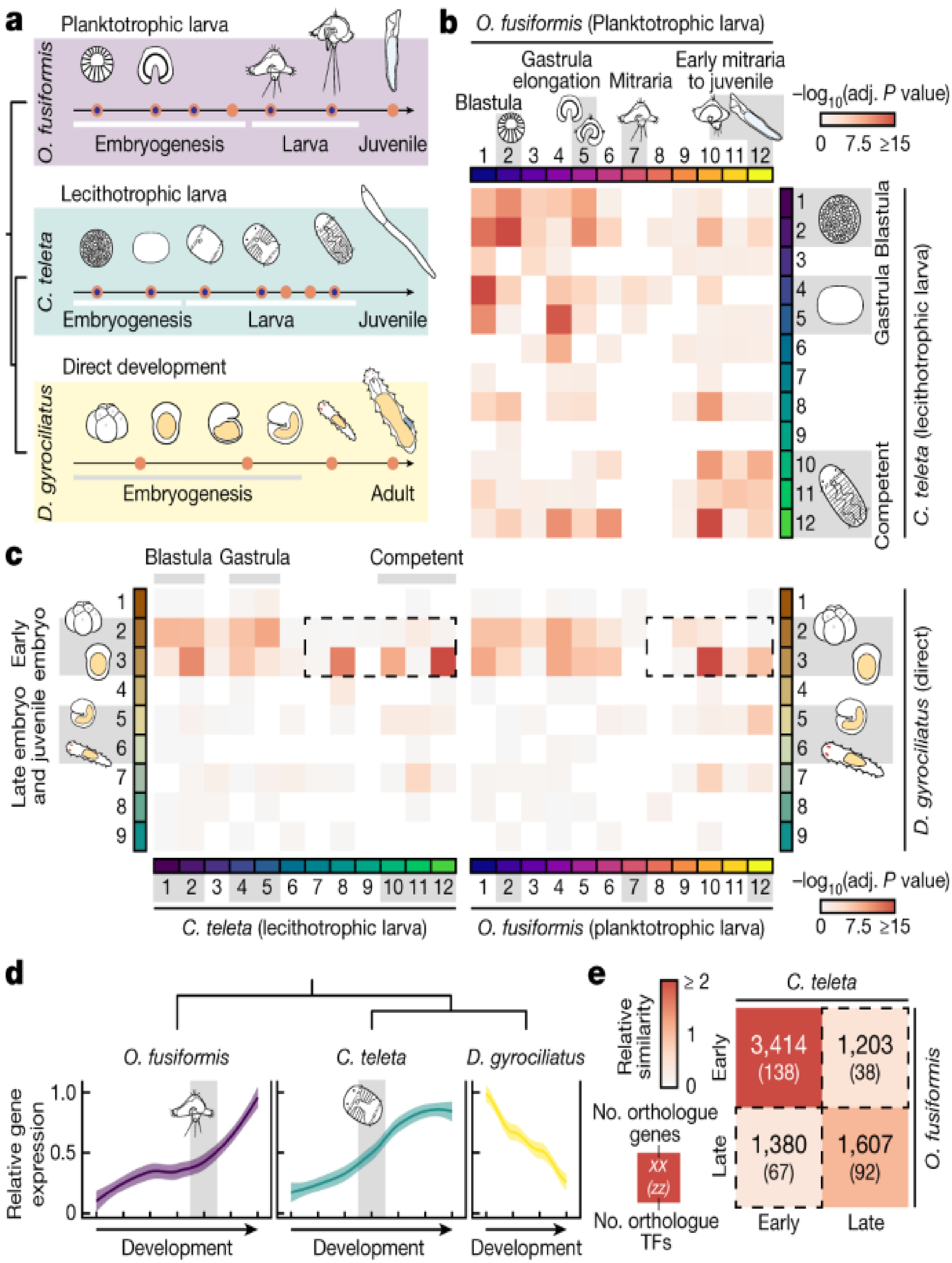
To investigate how larvae evolved in Annelida, we first generated a chromosome-scale reference assembly for the oweniid *O. fusiformis* (Fig. 1b, inset). The haploid assembly spans 505.8 Mb and has 12 chromosome-scale scaffolds (Supplementary Fig. 1). Almost half of the assembly (43.02%) consists of repeats (Extended Data Fig. 1b,c), and we annotated 26,966 protein-coding genes and 31,903 transcripts, which represent a nearly complete (97.5%) set of metazoan BUSCO genes (Supplementary Fig. 1). Gene family reconstruction and gene content analysis nested *O. fusiformis* with other non-annelid spiralians and taxa with slow-evolving genomes (Fig. 1b and Extended Data Fig. 1d,e). This result provides evidence that *O. fusiformis* has fewer gene family gains and losses and retains more ancestral metazoan orthogroups than other annelid taxa (Fig. 1c and Extended Data Fig. 1f,g). Indeed, *O. fusiformis* has a *chordin* orthologue, a bone morphogenetic protein inhibitor involved in dorsoventral patterning thought to be lost in annelids[27] and is asymmetrically expressed around the blastopore of the gastrula and larval mouth in *O. fusiformis* (Extended Data Fig. 2). Moreover, *O. fusiformis* has globally retained the ancestral bilaterian linkage, exhibiting chromosomal fusions that are present in molluscs and even nemerteans, and fewer lineage-specific chromosomal rearrangements than other annelids (Fig. 1d and Extended Data Fig. 1h,i). Therefore, *O.*

fusiformis shows a more complete gene repertoire and ancestral syntenic chromosomal organization than other annelids. Together with its phylogenetic position and conserved early embryogenesis[23,24], *O. fusiformis* is a key lineage to reconstruct the evolution of Annelida, and of Spiralia generally.

Heterochronies in gene expression

Next, we sought to identify transcriptomic changes that underpin the distinct life cycles in Annelida. We compared temporal series of embryonic, larval and competent and juvenile transcriptomes of *O. fusiformis* and *C. teleta*, two indirect developers with planktotrophic and lecithotrophic[28] larvae, respectively, and *D. gyrociliatus*, a direct developer[29,30] (Fig. 2a). Transcriptional dynamics during early embryogenesis were overall similar among these species (Supplementary Fig. 3). *C. teleta* and *D. gyrociliatus* showed increasing transcriptomic divergence with each other as they develop into adult stages; however, the maximal transcriptomic divergence between these annelids and *O. fusiformis* occurred at the mitraria stage (Extended Data Fig. 3a,b). Soft clustering of all expressed transcripts produced 12 distinct groups of temporally co-regulated genes in *O. fusiformis* and *C. teleta*, and 9 clusters in *D. gyrociliatus* (Extended Data Fig. 3c–e), which were expressed gradually along the life cycle of all three species. Only one cluster in each species showed a bimodal activation at early embryogenesis and in the competent larva (juvenile or adult forms), consistently involving genes enriched for core cellular processes (Extended Data Fig. 3f). Indeed, translation and metabolism predominated in clusters of early development in the three annelids, whereas cell communication and signalling, morphogenesis and organogenesis were enriched in later stages of development (Extended Data Fig. 3f). Therefore, regardless of the life cycle, transcriptional dynamics are generally conserved during annelid development, yet adults and the planktotrophic larva are the most transcriptionally distinct stages.

Fig. 2: Heterochronies in gene regulatory programmes underpin annelid life cycle diversification.



a, Experimental design of the comparative developmental RNA sequencing (RNA-seq) and assay for transposase-accessible chromatin with sequencing

(ATAC-seq) time courses. Orange circles highlight stages of *O. fusiformis*, *C. teleta* and *D. gyrociliatus* development sampled for bulk RNA-seq. Orange circles with a blue inner dot highlight developmental stages sampled for ATAC-seq. **b**,**c**, Similarity heatmaps showcasing the orthogroup overlap between the transcription factors contained in clusters of co-regulated genes obtained by soft *k*-means clustering between all three studied annelid taxa: 12 clusters for *O. fusiformis* and *C. teleta*, and 9 clusters for *D. gyrociliatus*. Time points associated to key clusters are shown for all three species. Dotted black lines in **c** highlight the distinct timing expression differences of a significant number of transcription factors shifted from post-larval expression in indirect developers to early embryogenesis in *D. gyrociliatus*. *P* values were derived from upper-tail hypergeometric tests and Benjamini–Hochberg-adjusted (adj. *P* value). **d**, Average expression dynamics of the 28 single-copy orthologue transcription factors shifted from late expression in both *O. fusiformis* and *C. teleta* to early expression in *D. gyrociliatus*. Curves are locally estimated scatterplot smoothings, coloured shaded areas represent standard error of the mean. **e**, Heatmap of relative similarity based on whole-genome orthogroup overlap analysis by quadrants between pre-larval (early) and post-larval (late) clusters in *O. fusiformis* and *C. teleta*. Dotted black lines denote the groups of genes and transcription factors under heterochronies between both species.

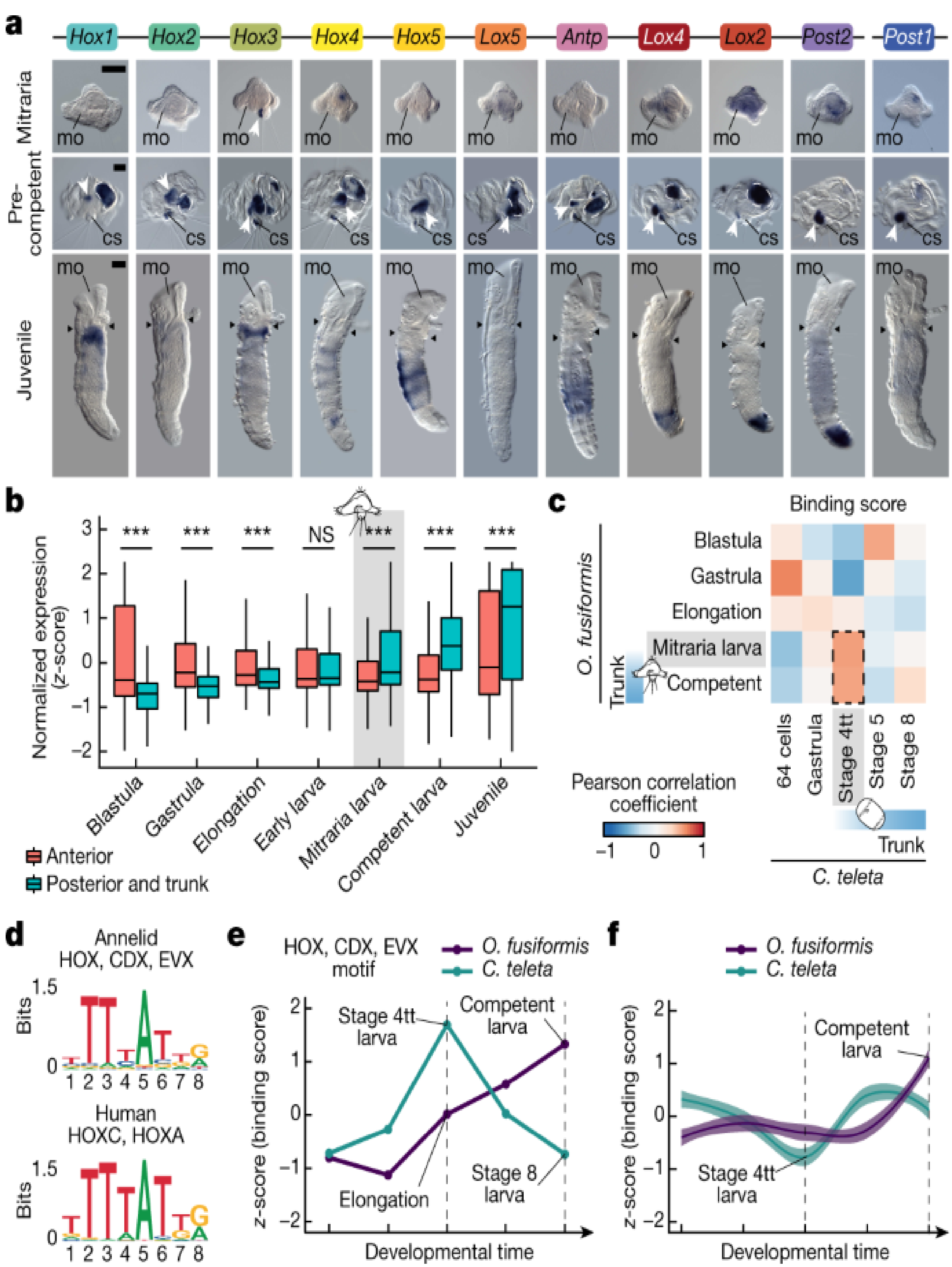
To identify the genes that mediate the transcriptional differences at larval and adult stages, we performed pairwise inter-species comparisons of gene and transcription factor composition among clusters of temporally co-regulated genes (Fig. 2b,c and Extended Data Fig. 4a,b). Early clusters followed by late clusters were the most conserved in the three comparisons when all genes were considered (Extended Data Fig. 4c,d). However, transcription factors used in post-larval stages in indirect development were consistently shifted towards early embryogenesis in direct development (Fig. 2c and Extended Data Fig. 4c,e). In both *O. fusiformis* and *C. teleta*, this shift involved 28 transcription factors that function in various developmental processes, from nervous system (for example, *pax6* (ref. [31])) and mesoderm (for example, *foxF* (ref. [26])) formation to axial patterning (for example, *Hox1* and *Hox4* (ref. [32])) (Supplementary Fig. 12). Notably, the overall expression of these 28 genes was also temporally shifted between indirect developing annelids, with the maximum level of expression occurring earlier

in *C. teleta* than in *O. fusiformis* (Fig. 2d). Additionally, 2,583 genes also exhibited temporal shifts between the larvae of *O. fusiformis* and *C. teleta* (Fig. 2e), including 105 transcription factors, but mostly enzymes and structural genes that probably reflect the different biology of these two larvae (Extended Data Fig. 4f,g and Supplementary Figs. 13–16). Therefore, temporal shifts (that is, heterochronies) in the use of shared genetic programmes and regulatory genes correlate with and might account for life cycle and larval differences in Annelida.

Different timings of trunk development

Homeodomain transcription factors were the largest class among the 28 transcription factors with temporal expression shifts between direct and indirect developing annelids (Supplementary Fig. 12). Indeed, homeodomain genes were enriched in the competent larva in *O. fusiformis* but were prevalent from stage 5 larva onwards in *C. teleta* (Extended Data Fig. 4h). Accordingly, Hox genes, which regionalize the bilaterian trunk along the anteroposterior axis[33], were strongly upregulated in the competent mitraria larva (Extended Data Fig. 5a,b). *O. fusiformis* had a conserved complement of 11 Hox genes—similar to *C. teleta*[32]—arranged as a compact, ordered cluster in chromosome 1, except for *Post1*, which was located downstream of this chromosome (Extended Data Fig. 5c,d). *C. teleta* and *D. gyrociliatus* started expressing Hox genes along their trunks[30,32] during or soon after gastrulation (Extended Data Fig. 5e). *O. fusiformis*, however, did not express Hox genes during embryogenesis but in the trunk rudiment during larval growth, already in an anteroposterior staggered pattern, as later observed in the juvenile (Fig. 3a and Extended Data Fig. 5e–h). This late activation of Hox genes is not specific to *O. fusiformis*, as it also occurs for most Hox genes in the planktotrophic trochophore of the echiuran annelid *Urechis unicinctus*[34] (Extended Data Fig. 5e). Therefore, the spatially collinear Hox code along the trunk is established at distinct developmental stages depending on the life cycle mode in Annelida.

Fig. 3: Trunk development is delayed to pre-metamorphosis in *O. fusiformis*.



a, Representative images from three independent analyses of Hox gene expression by whole-mount in situ hybridization at the mitraria larva, pre-competent larva, and juvenile stages of *O. fusiformis*. Only *Hox3* is

expressed at the mitraria stage (white arrowhead). Hox genes show spatial collinearity along the anteroposterior axis at the developing trunk of the pre-competent larva (white arrowheads) and in the juvenile. Dotted lines in the competent larva panels indicate background from the midgut. Black arrowheads in the juvenile panels indicate head to trunk boundary. cs, chaetal sack; mg, mid gut. Scale bar, 50 µm (larval stages) or 100 µm (juvenile stage). **b**, Average expression dynamics of anterior ($n = 1,655$), and posterior and trunk genes ($n = 407$) expressed in corresponding adult tissues during *O. fusiformis* development. *P* values were derived from two-tailed Student's *t*-tests and adjusted using the Bonferroni method for multiple testing correction. ***, $P < 0.001$; NS, not significant. Centre lines in boxplots are the median, box is the interquartile range (IQR), and whiskers are the first or third quartile ± 1.5× IQR. **c**, Correlation matrices of transcription factor binding score (TFBSs). The dotted black line highlights the high TFBS correlation and heterochrony between the mitraria and competent larvae of *O. fusiformis* and the stage 4tt larva of *C. teleta*. **d**, Sequence logo of the annelid archetype (top) shows substantial similarity to the human homologue (bottom). **e**, TFBS dynamics for the annelid HOX, CDX, and EVX motif during *O. fusiformis* and *C. teleta* development. **f**, Average TFBS dynamics of all motifs in the peaks of the Hox cluster. Curves are locally estimated scatterplot smoothings, coloured shaded areas represent standard error of the mean.

To determine whether the difference in timings of trunk patterning is limited to the expression of Hox genes, we used tissue-specific adult transcriptomes to define a set of 1,655 anterior and 407 posterior and trunk genes in *O. fusiformis* (Extended Data Fig. 6a–d). Anterior genes were significantly more expressed during embryogenesis, whereas posterior and trunk genes were upregulated at the mitraria stage and significantly outweighed the expression dynamics of anterior genes from that stage onwards (Fig. 3b and Extended Data Fig. 6e,f). Moreover, anterior, trunk and posterior genes with spatially resolved expression followed different temporal dynamics in *O. fusiformis*, *C. teleta* and *D. gyrociliatus*. In *O. fusiformis*, trunk[25] and posterior[24,26] genes were concentrated in a small ventral area and around the anal opening of the larva and increased in spatial range and expression levels as the trunk formed (Extended Data Fig. 6g,h). By contrast, anterior genes[26,35] patterned most of the mitraria, and their expression remained

stable during development (Extended Data Fig. 6g,h). Posterior and anterior genes followed similar dynamics in *C. teleta*, and trunk genes were upregulated already post-gastrula in both *C. teleta* and *D. gyrociliatus* (Extended Data Fig. 6i–l). Therefore, trunk development, which initially occurs from lateral growth of the trunk rudiment[12,28], is deferred to pre-metamorphic stages in planktotrophic annelid trochophores compared with annelids with lecithotrophic larvae and direct developers.

Heterochronies in Hox regulation

To investigate the genomic regulatory basis for the heterochronies in trunk development among annelid larvae, we profiled open chromatin regions at five equivalent developmental stages in *O. fusiformis* and *C. teleta* (Fig. 2a). This analysis identified 63,726 and 44,368 consensus regulatory regions, respectively. In both species, open chromatin was more abundant within gene bodies (Extended Data Fig. 7a). There was, however, a general increase in promoter peaks in *O. fusiformis* and distant intergenic regulatory elements in both species during development (Extended Data Fig. 7b). Moreover, the largest changes in peak accessibility occurred in the mitraria in *O. fusiformis* and stage 5 larva in *C. teleta* (Supplementary Fig. 18). In *O. fusiformis*, most regulatory regions acted before the start of trunk formation, whereas the numbers of accessible regions with a maximum of accessibility before and after the onset of trunk development were comparable in *C. teleta* (Extended Data Fig. 7c). Accordingly, the regulation of genes involved in morphogenesis and organogenesis, as well as neurogenesis, was concentrated in late clusters in *O. fusiformis* but unfolded more continuously in *C. teleta* (Supplementary Fig. 23). Therefore, different dynamics of chromatin accessibility occur during development and larva formation in these two annelids.

To investigate the regulatory programmes controlling larva development in *O. fusiformis* and *C. teleta*, we predicted transcription factor-binding motifs on peaks obtained from ATAC-seq data. This analysis identified 33 motifs common to both species that were strongly assigned to a known transcription factor class (Supplementary Fig. 29). Notably, the binding dynamics of these 33 motifs revealed a temporal shift in regulatory motifs acting between the mitraria and competent larva in *O. fusiformis* to the early

post-gastrula (stage 4tt) larva of *C. teleta* (Fig. 3c and Extended Data Fig. 7d–f). Seven motifs followed this pattern (Extended Data Fig. 7g and Supplementary Fig. 29), including one with high similarity to the human HOX, CDX and EVX motif archetype (Fig. 3d,e) that is overrepresented and upregulated on the basis of its binding score at the competent stage in *O. fusiformis* (Extended Data Fig. 7h and Supplementary Fig. 30). Indeed, motif-binding dynamics in regulatory elements assigned to Hox genes supported a change in global regulation of the Hox cluster at the competent and early larval stages in *O. fusiformis* and *C. teleta*, respectively (Fig. 3f and Supplementary Fig. 31), which mirrored the transcriptional onset of these genes and the start of trunk development in the two species[32]. Motifs assigned to NKX and GATA factors, which are expressed in the developing trunk in both species[25,36], were among the most abundant bound motifs in the Hox cluster in both species (Extended Data Fig. 7i). However, only 39 one-to-one orthologues with bound HOX, CDX and EVX motifs at the maximum of motif binding were common to *O. fusiformis* and *C. teleta* (Extended Data Fig. 7j). Therefore, different regulatory dynamics of the Hox cluster—possibly triggered by a reduced common set of upstream regulators—underpin temporal variability in Hox activity and downstream targets. These shifts probably promoted the developmental and morphological differences in trunk formation between planktotrophic and lecithotrophic annelid larvae.

Different dynamics of new genes

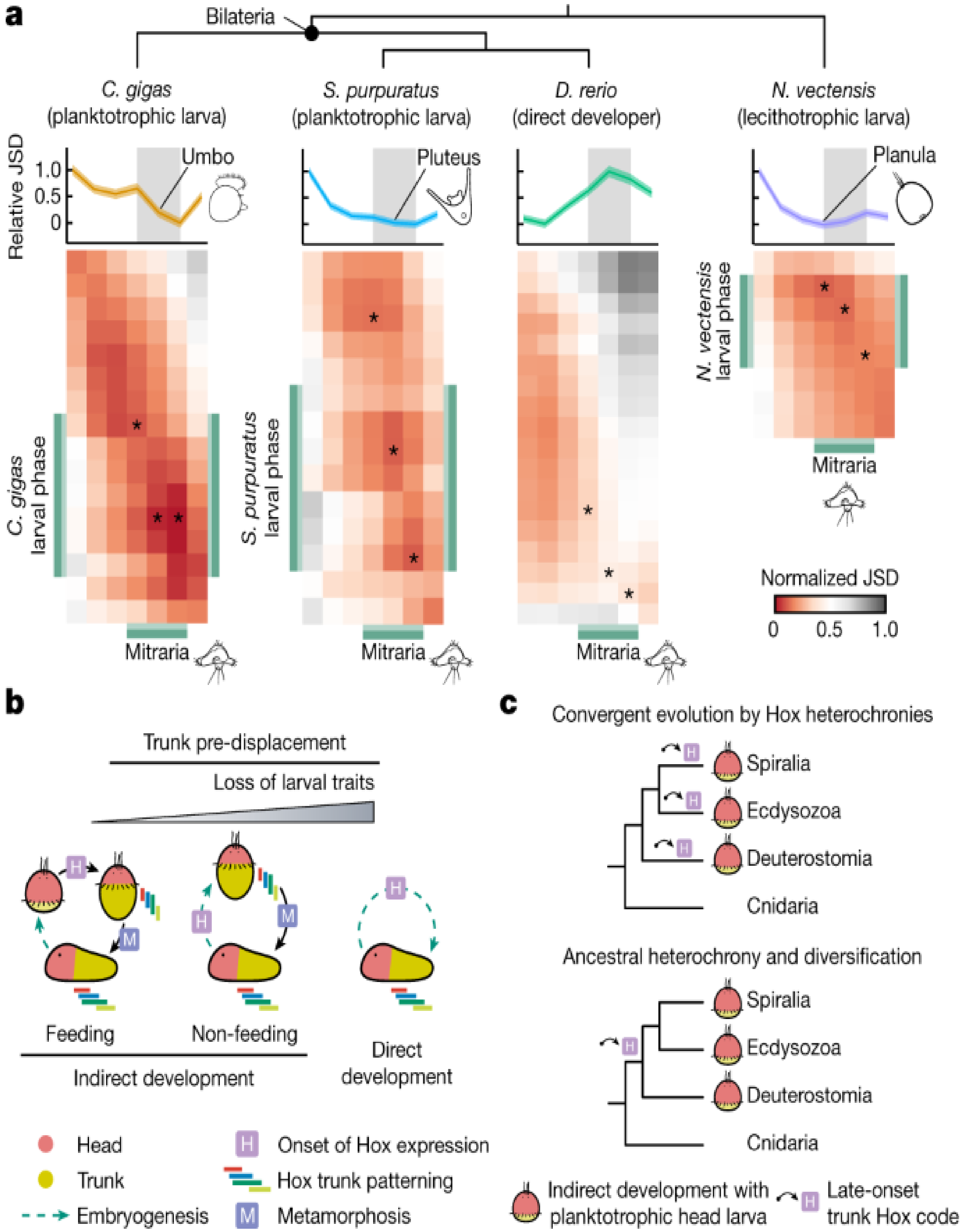
New, species-specific genes, which account for a significant proportion of some larval transcriptomes[6,37], could also contribute to and explain the transcriptomic differences among annelid larvae. In *O. fusiformis*, *C. teleta* and *D. gyrociliatus*, genes of metazoan and pre-metazoan origin tended to peak, dominate and be enriched at early development, whereas younger genes were more highly expressed in competent and juvenile stages (Extended Data Fig. 8a–e). By contrast, species-specific genes followed lineage-specific dynamics (Supplementary Fig. 32). These genes, for instance, were more expressed in the juveniles of *O. fusiformis* and *D. gyrociliatus*, but in the blastula and gastrula of *C. teleta* (and to some extent also at the blastula stage in *O. fusiformis*; Extended Data Fig. 8a,c,d). Species-specific genes were only enriched and over-represented at larval

stages in *C. teleta* (Extended Data Fig. 8f–h). Therefore, genes of different evolutionary origins contribute to the development of annelid larvae. This result suggests that the increased use of new genes in some lophotrochozoan larvae[6,37] might be due to the evolution of lineage-specific larval traits.

Similarities between bilaterian larvae

To assess whether the transcriptional dynamics found in annelids are also observed in other metazoans, we extended our comparative transcriptomic approach to nine other animal lineages. In relative terms, global transcriptional dynamics between *O. fusiformis* and other animals tended to be more dissimilar at early development than at juvenile and adult stages (Fig. 4a and Extended Data Figs. 9a,b and 10a). The exception was the direct developer *Danio rerio*, for which the mitraria larva was the most dissimilar stage (Fig. 4a). This was also the case when comparing *O. fusiformis* with the direct-developing annelid *D. gyrociliatus* (Extended Data Fig. 3b). Notably, *O. fusiformis* shared maximal transcriptomic similarities during larval phases with bilaterian species with planktotrophic ciliated larvae and even cnidarian planulae (Fig. 4a and Extended Data Fig. 9a–e). Genes involved in core cellular processes directly contributed to these similarities, which probably reflects common structural and ecological needs of metazoan larvae (Extended Data Fig. 9f,g). However, transcription factor expression levels were also maximally similar between those species at larval phases (Extended Data Fig. 9a,b,e). Therefore, adult development is generally more similar[9] than early embryogenesis across major animal lineages, but phylogenetically distant animal larvae also exhibit unexpected genome-wide transcriptional—and potentially regulative—similarities.

Fig. 4: The evolution of life cycles in Annelida and Bilateria.



a, Heatmaps of pairwise normalized Jensen–Shannon divergence (JSD) between *O. fusiformis* and *Crassostrea gigas*, *Strongylocentrotus purpuratus*, *D. rerio* and *Nematostella vectensis*. Asterisks indicate the stages of minimal JSD of each species to the larval phase of *O. fusiformis*.

Larval phases are highlighted in green. Average relative JSD of the stages of minimal divergence to each *O. fusiformis* stage is shown on top of each heatmap. Confidence intervals represent the standard deviation from 250 bootstrap resamplings of the orthologue sets. See Extended Data Fig. 9a and Extended data Fig. 10a for fully labelled heatmaps. **b**, Schematics of the three main types of life cycles and the timing of Hox gene expression in bilaterians. Compared to indirect development with feeding larvae, lineages with non-feeding larvae and direct development pre-displace (that is, initiate earlier) trunk differentiation and Hox gene expression. Larval organs are reduced in non-feeding larvae and absent in direct development. **c**, Proposed alternative scenarios for the evolution of maximal indirect development with head larvae in Bilateria. Top, head larvae evolved convergently by repeatedly shifting trunk development (as seen by Hox gene expression) to pre-metamorphic stages. Bottom, head and trunk development were ancestrally temporally decoupled, which could have facilitated the evolution of head larvae in different bilaterian lineages.

Discussion

Our study provides a perspective on life cycle evolution in Bilateria. The planktotrophic larva of *O. fusiformis* defers trunk differentiation to late pre-metamorphic stages and largely develops from anterior ectodermal domains. This occurs in other feeding annelid larvae[38] (Extended Data Fig. 5f), and probably in Chaetopteriformia[39,40], and thus the late differentiation of the adult trunk might be an ancestral trait to Annelida (Extended Data Fig. 10b). Delaying trunk development to post-larval stages also occurs in phylogenetically distant clades within Spiralia[16,17], Ecdysozoa[14,41] and Deuterostomia[15,42,43], the larvae of which are generally referred to as head larvae[13,14]. By contrast, non-feeding larvae[32,44] and direct developers[30] in both Annelida and other bilaterian taxa[45,46] start to pattern their trunks with or immediately after the onset of anterior or head patterning, which always takes place before gastrulation in bilaterians[47,48]. Therefore, heterochronies in trunk development correlate with, and possibly account for, the evolution of different life cycles in animals (Fig. 4b). This differs from previously proposed mechanisms to explain the origins of animal life cycles, namely

co-option of adult genes into larval-specific regulatory programmes[9,10] and independent evolution of adult gene regulatory modules[2,49].

Bilaterian head larvae could be lineage-specific innovations associated with the evolution of maximal indirect development[13,14,16] that evolved convergently by delaying trunk differentiation and Hox patterning (Fig. 4c). The similarities in larval molecular patterns[5,15,16] would then reflect ancient gene regulatory modules that were independently co-opted to develop analogous cell types and larval organs. Alternatively, the post-embryonic onset of trunk differentiation and Hox expression might be the most parsimonious ancestral state for Bilateria (Extended Data Fig. 10c,d and Supplementary Table 93). This could have facilitated the evolution of larvae, which would then originally share anterior genetic modules for their development (Fig. 4c). Regardless of the scenario and despite their limitations, our datasets highlight the importance of heterochronic changes for the diversification of bilaterian life cycles. The data also uncover a reduced set of candidate genes and regulatory motifs that might influence life cycle differences in Annelida and perhaps even Bilateria. In the future, comparative functional studies of these and other genes will reveal how temporal changes in gene expression and regulation have shaped the evolution of animal larvae and adults.

Methods

Adult culture, spawning and in vitro fertilization

Sexually mature *O. fusiformis* adults were collected from subtidal waters near the Station Biologique de Roscoff and cultured in the laboratory as previously described[23]. In vitro fertilization and collection of embryonic and larval stages were performed as previously described[23]. *C. teleta* Blake, Grassle & Eckelbarger, 2009 was cultured, grown and sifted, and its embryos and larvae were collected following established protocols[28]. *Magelona* spp. were collected in muddy sand from the intertidal of Berwick-upon-Tweed, Northumberland, NE England (around 55° 46′00.4″ N, 1° 59′04.5″ W) and kept initially in aquaria at the Museum Wales before their

transfer to Queen Mary University of London, where they were kept in aquaria with artificial sea water.

Genome size measurements

To estimate the haploid DNA nuclear content of *O. fusiformis*, we used a flow cytometer Partex CyFlow Space fitted with a Cobalt Samba green laser (532 nm, 100 mW) and the built-in software FloMax (v.2.82) as described for the annelid *D. gyrociliatus*[23], with adult individuals of *Drosophila melanogaster* as reference. Additionally, we used Jellyfish (v.2.3)[50] to count and generate a 31-mer histogram from adaptor-cleaned, short-read Illumina reads (see section below) and GenomeScope (v.2.0)[51] to obtain an in silico estimation of the genome size and heterozygosity of *O. fusiformis*.

Genome sequencing, assembly and quality check

Ultra-high molecular weight (UHMW) gDNA was extracted following the Bionano genomics IrysPrep agar-based, animal tissue protocol using sperm from a single *O. fusiformis* male. UHMW gDNA was cleaned up using a salt–chloroform wash following PacBio's recommendations before long-read sequencing using PacBio (v.3.0) chemistry at the University of California Berkeley. A total of 16 SMRT cells of PacBio Sequel were used for sequencing with 600 min movie time, producing a total of 170.07 Gb of data (10.72 million reads, N50 read length between 25.75 kb and 30.75 kb). In addition, we used UHMW gDNA of that same individual to generate a 10x Genomics linked reads library, which we sequenced in an Illumina HiSeq4000 at the Okinawa Institute of Science and Technology to produce 28.62 Gb of data (141.66 million read pairs). PacBio reads were assembled with CANU (v.8.3rc2)[52] assuming 'batOptions="−dg 3 −db 3 −dr 1 −ca 500 −cp 50' and 'correctedErrorRate = 0.065'. Pacbio reads were remapped using pbalign (v.0.3.2) and the assembly polished once using Arrow (genomicconsensus, v.2.3.2). Then Illumina paired-end reads generated with the 10x Genomics linked reads were extracted, remapped using bwa mem (v.0.7.17)[53] and used for polishing with Racon (v.1.16)[54]. Bionano Genomics optical mapping data were used to scaffold the PacBio-based assembly, which was de-haploidized with purge_haplotigs (v.1.0.4)[55] setting cut-off values at 35, 85 and 70× coverage to reconstruct a high-quality

haploid reference assembly. HiC-based chromosome scaffolding was performed as described below. Merqury (v.1.1)[56] and BUSCO (v.5)[57] were used to assess genome completeness and to evaluate the quality of the assembly (Supplementary Fig. 1).

Transcriptome sequencing

Fourteen samples spanning key developmental time points of the *O. fusiformis* life cycle, including active oocyte, zygote, 2-cell, 4-cell and 8-cell stages, 3 h post-fertilization (h.p.f.), 4 h.p.f., coeloblastula (5 h.p.f.), gastrula (9 h.p.f.), axial elongation (13 h.p.f.), early larva (18 h.p.f.), mitraria larva (27 h.p.f.), pre-metamorphic competent larva (3 weeks post-fertilization) and post-metamorphic juvenile were collected in duplicates (except for the latter), flash frozen in liquid nitrogen and stored at –80 °C for total RNA extraction. Samples within replicates were paired, with each one containing around 300 embryos or 150 larvae coming from the same in vitro fertilization process. Nine further samples from adult tissues and body regions (blood vessel, body wall, midgut, prostomium, head, ovary, retractor muscle, tail and testes) were also collected as described above. Likewise, an additional five samples spanning post-cleavage time points of *C. teleta*, including 64 cells and gastrula stages, and stage 4tt, stage 5 and stage 7 larval stages, were collected in duplicates. Total RNA was isolated using a Monarch Total RNA Miniprep kit (New England Biolabs) following the supplier's recommendations. Total RNA samples from developmental stages from both *O. fusiformis* and *C. teleta* were used to prepare strand-specific mRNA Illumina libraries that were sequenced at the Oxford Genomics Centre (University of Oxford, UK) over three lanes of an Illumina NovaSeq6000 system in 2 × 150 bp mode to a depth of around 50 million reads (Supplementary Tables 13 and 16). Adult tissue samples were sequenced at BGI on a BGISeq-500 platform in 2 × 100 bp mode to a depth of about 25 million reads (Supplementary Table 49).

Annotation of repeats and transposable elements

RepeatModeler (v.2.0.1)[58] and RepBase were used to construct a de novo repeat library for *O. fusiformis*, which was then filtered for bona fide genes using the predicted proteome of *C. teleta*. In brief, we used DIAMOND

(v.0.9.22)[59] with an *e*-value cut-off of 1×10^{-10} to identify sequences in the de novo repeat library with significant similarity to protein-coding genes in *C. teleta* that are not transposable elements (TEs). Sequences with a significant hit were manually inspected to verify they were not TEs; if they were, they were manually removed from the de novo repeat library. The filtered consensus repeat predictions were then used to annotate the genome assembly of *O. fusiformis* with RepeatMasker open-4.0. We next used LTR_finder (v.1.07)[60], a structural search algorithm, to identify and annotate long tandem repeats (LTRs). Finally, we generated a consensus set of repeats by merging RepeatMasker and LTR_finder predictions with RepeatCraft[61], using default parameters but a maximum LTR size of 25 kb (as derived from the LTR_finder annotation) (Supplementary Table 1). The general feature format (GFF) and fasta files with the annotation of TEs and repeats are available in the GitHub repository (see Data availability section).

Gene prediction and functional annotation

We used SAMtools (v.1.9)[62] and the annotation of repeats to soft mask *O. fusiformis* genome assembly before gene prediction. We then mapped all embryonic and adult transcriptomes and a publicly available dataset[63] (Sequence Read Archive (SRA) identifier: SRR1222288) with STAR (v.2.5.3a)[64] after removing low-quality read pairs and read pairs containing Illumina sequencing adapters with trimmomatic (v.0.39)[65]. StringTie (v.1.3.6)[66] was used to convert STAR alignments into gene transfer format (GTF) files and Portcullis (v.1.1.2)[67] to generate a curated set of splice junctions. Additionally, we generated de novo transcriptome assemblies for all samples using Trinity (v.2.5.1)[68] with default parameters, which were thereafter mapped to the soft-masked assembly with GMAP (v.2020-04-08)[69]. We then ran the default Mikado (v.2.1) pipeline[70] to merge all transcriptomic evidence and reliable splice junctions into a single set of best-supported transcripts and gene models. From this merged dataset, we filtered full-length, non-redundant transcripts with a BLAST hit on at least 50% of their length and at least two exons to obtain a gene set that we used to train Augustus (v.3.2.3)[71]. Simultaneously, we used the Mikado gene annotation and Portcullis splice junctions to generate confident sets of exon and intron hints, respectively. We also ran Exonerate (v.2.4.0)[72] to generate spliced

alignments of the proteome of *C. teleta* proteome on *O. fusiformis* soft-masked genome assembly to obtain further gene hints. We then merged all exon and intron hints into a single dataset, which we passed into Augustus (v.3.2.3)[71] for ab initio gene prediction. Finally, PASA (v.2.3.3)[73] was used to combine RNA-seq and ab initio gene models into a final gene set, from which spurious predictions with in-frame stop codons (228 gene models), predictions that overlapped with repeats (5,779 gene models) and that had high similarity to TEs in the RepeatPeps.lib database (2,450 models) were removed. This filtered gene set included 26,966 genes, encompassing 31,903 different transcripts (Supplementary Fig. 1). To assess the completeness of this annotation, we ran BUSCO (v.5)[57] in proteome mode, which resulted in 97.7% of the core genes present. Moreover, 31,678 out of the 31,903 (99.29%) filtered transcripts were supported by RNA-seq data, and 80.69% of the transcripts had a significant BLAST match (*e*-value cut-off < 0.001) to a previously annotated annelid gene (database containing non-redundant proteomes of the high-quality annelid genomes of *C. teleta*, *D. gyrociliatus*, *Eisenia andrei*, *Lamellibrachia luymesi*, *Paraescarpia echinospica*, *Riftia pachyptila* and *Streblospio benedicti*). A similar functional annotation approach was followed to re-annotate the genome of *C. teleta* with the new RNA-seq data, using as starting assembly the soft-masked version available at Ensembl Metazoa. This resulted in 41,221 transcripts, 39,814 of which had RNA-seq support (96.59%). Additionally, 80.47% of the transcripts had a significant BLAST match (*e*-value cut-off < 0.001) to other well-annotated annelid genomes (see above).

Protein homologies for the filtered transcripts of *O. fusiformis* and *C. teleta* were annotated using BLAST (v.2.2.31+)[74] with the UniProt/SwissProt database provided with Trinotate (v.3.0)[75]. We used HMMER (v.2.3.2)[76] to identify protein domains using Trinotate's PFAM-A database and signalP (v.4.1)[77] to predict signal peptides. These functional annotations were integrated into a Trinotate database, which retrieved Gene Ontology (GO), eggNOG and Kyoto Encyclopedia of Genes and Genomes (KEGG) terms for each transcript. In addition, we ran the PANTHER HMM scoring tool to assign a PantherDB[78] orthology identifier to each transcript. In total, we retrieved a functional annotation for 22,516 transcripts (63.86%). Functional annotation reports are provided in the GitHub repository (see Data Availability section).

Chromosome-scale scaffolding

Sperm from a single *O. fusiformis* worm and an entire sexually mature male were used as input material to construct two Omni-C Dovetail libraries following the manufacturer's recommendations for marine invertebrates. These libraries were sequenced in an Illumina NovaSeq6000 at the Okinawa Institute of Science and Technology to a depth of 229 and 247 million reads. HiC reads were processed using the Juicer pipeline (r.e0d1bb7)[79] to generate a list of curated contracts ('merged no dups') that was subsequently used to scaffold the assembly using 3d-dna (v.180419)[80]. The resulting assembly and contact map were visually inspected and curated using Juicebox (v.1.11.08)[79], and adjustments were submitted for a subsequent run of optimization using 3d-dna. Finally, repeats and TEs were re-annotated in this chromosome-scale assembly as described above, and the annotation obtained for the PacBio-based assembly was lifted over with Liftoff (v.1.6.1)[81] (Supplementary Fig. 1). All gene models but two were successfully re-annotated in the chromosome-scale assembly.

Gene family evolution analyses

We used the AGAT suite of scripts to generate non-redundant proteomes with only the longest isoform for a set of 21 metazoan proteomes (Supplementary Table 2). To reconstruct gene families, we used OrthoFinder (v.2.2.7)[82] using MMSeqs2 (ref. [83]) to calculate sequence similarity scores and an inflation value of 2. OrthoFinder gene families were parsed and mapped onto a reference species phylogeny to infer gene family gains and losses at different nodes and tips using the ETE 3 library[84], as well as to estimate the node of origin for each gene family. Gene expansions were computed for each species using a hypergeometric test against the median gene number per species for a given family using previously published code[30] (Supplementary Tables 3–7). Principal component analysis was performed on the orthogroups matrix by metazoan lineage, given that orthogroups were present in at least three of the 22 analysed species, to eliminate taxonomically restricted genes. All single copy orthologue files derived from this analysis used throughout the study are available in the GitHub repository (see Data Availability section).

Macrosynteny analyses

Single-copy orthologues obtained using the mutual best hit approach generated using MMseqs2 (ref. [83]) using the annotations of *Branchiostoma floridae*[85], *P. maximus*[86], *S. benedictii*[87] and *Lineus longissimus*[88,89] were used to generate Oxford synteny plots comparing sequentially indexed orthologue positions. Plotting order was determined by hierarchical clustering of the shared orthologue content using the complete linkage method as originally proposed. Comparison of the karyotype of all four species was performed using the Rideogram package by colouring pairwise orthologues according to the ALG assignment in comparisons with *P. maximus* and *B. floridae*.

Evolutionary analysis of *chordin* in annelids

The identification of *chordin* (*chrd*) and *chordin-like* (*chrdl*) genes in *O. fusiformis* was based on the genome functional annotation (see above). To mine *chrd* orthologues, 81 annelid transcriptomic datasets were downloaded from the SRA (Supplementary Table 8) and assembled using Trinity (v.2.5.1)[68] to create BLAST local nucleotide databases. We also created a nucleotide database for *C. teleta* using its annotated genome[90] (European Nucleotide Archive (ENA) accession number: GCA_000328365.1). Human and *O. fusiformis* CHRD proteins were used as queries to find *chrd* orthologues following the mutual best hit approach (*e*-value $\leq 10^{-3}$), obtaining 103 distinct candidate *chrd* transcripts that were then translated (Supplementary Table 9). A single candidate CHRD protein for *Themiste lageniformis* (M. J. Boyle, unpublished data) was included ad hoc at this step. In addition, 15 curated CHRD and CHRDL protein sequences (and an outgroup) were obtained from various sources (Supplementary Table 10) and aligned together with *O. fusiformis* CHRD and CHRDL sequences in MAFFT (v.7)[91] with the G-INS-I iterative refinement method and default scoring parameters. From this mother alignment, further daughter alignments were obtained using “mafft --addfragments”[92], the accurate “--multipair” method, and default scoring parameters. For orthology assignment, two phylogenetic analyses were performed on selected candidate sequences, which included the longest isoform for each species–gene combination, given that it included a 10-residue or longer properly

aligned fragment in either the CHRD domains or the von Willebrand factor type C (VWFC) domains. vWFC and CHRD domains were trimmed and concatenated using domain boundaries defined by ProSITE domain annotation for the human chordin precursor protein (UniProt: Q9H2X0). Either all domains or the VWFC domains only were used for phylogenetic inference (Extended Data Fig. 2c,d and Supplementary Tables 11 and 12) with a WAG amino acid replacement matrix[93] to account for transition rates, the FreeRate heterogeneity model (R4)[94] to describe sites evolution rates, and an optimization of amino acid frequencies using maximum likelihood using IQ-TREE (v.2.0.3)[95]. 1,000 ultrafast bootstraps[96] were used to extract branch support values. Bayesian reconstructions in MrBayes (v.3.2.7a)[97] were also performed using the same WAG matrix but substituting the R4 model for the discrete gamma model[98], with 4 rate categories (G4). All trees were composed in FigTree (v.1.4.4). Alignment files are available in the GitHub repository (see Data availability section).

Gene expression profiling

We profiled gene expression dynamics from blastula to juvenile stages for *O. fusiformis*, from 64-cell to competent larva stages for *C. teleta* (Supplementary Fig. 2), from early development to female adult stages for *D. gyrociliatus*, and across the 9 adult tissues samples of *O. fusiformis*. Sequencing adaptors were removed from raw reads using trimmomatic (v.0.39)[65]. Cleaned reads were pseudo-aligned to the filtered gene models using kallisto (v.0.46.2)[99], and genes with an expression level above an empirically defined threshold of 2 transcripts per million (TPM) were deemed expressed. For each species, the DESeq2 (v.1.30.1) package[100] was used to normalize read counts across developmental stages (Supplementary Tables 13–21) and adult tissues (Supplementary Tables 49–51) and to perform pairwise differential gene expression analyses between consecutive developmental stages. *P* values were adjusted using the Benjamini–Hochberg method for multiple testing correction. We defined a gene as significantly upregulated for a $\log_2$(fold-change) (LFC) > 1 or downregulated for a LFC < 1, given an adjusted *P* value < 0.05. Principal component analyses were performed on the variance stabilizing-transformed matrices of the normalized DESeq2 matrices. For the *O. fusiformis* adult

tissues samples, genes specifically expressed (TPM > 2) only in both the head and head plus two anterior-most segment samples were classified as adult anterior genes, and those expressed only in both the tail and the body wall were classified as adult trunk and posterior genes (Supplementary Tables 52 and 53). For all three annelid taxa, anterior, trunk and posterior markers were defined as genes for which their spatial expression pattern has been validated through in situ hybridization in the literature (Supplementary Tables 54–56). TPM and DESeq2 gene expression matrices of developmental and adult tissue samples are also available in the GitHub repository (see Data availability section).

Gene clustering and co-expression network analyses

Transcripts were clustered according to their normalized DESeq2 expression dynamics through soft *k*-means clustering (or soft clustering) using the mfuzz (v.2.52) package[101] (Supplementary Tables 23–26). Out of the total number of transcripts, we discarded those that were not expressed at any developmental stage (225 out of 31,903 for *O. fusiformis*, 1,407 out of 41,221 for *C. teleta*, and 200 out of 17,388 for *D. gyrociliatus*). We then determined an optimal number of 12 clusters (*O. fusiformis* and *C. teleta*) and 9 clusters (*D. gyrociliatus*) for our datasets by applying the elbow method to the minimum centroid distance as a function of the number of clusters. For construction of the gene co-expression networks for *O. fusiformis* and *C. teleta*, we used the WGCNA package (v.1.70-3)[102]. All transcripts expressed at any developmental stage were used to build a signed network with a minimum module size of 300 genes and an optimized soft-thresholding power of 16 and 8 for *O. fusiformis* and *C. teleta*, respectively. Block-wise network construction returned 15 gene modules for *O. fusiformis*, from which 1 module was dropped owing to poor intramodular connectivity, and 19 gene modules for *C. teleta* (Supplementary Tables 23 and 24). The remaining 14 gene modules of *O. fusiformis* (A–N) and 19 gene modules of *C. teleta* (A–O, W–Z) were labelled with distinct colours, with unassigned genes labelled in grey. Random subsets consisting of the nodes and edges of 30% of the transcripts were fed into Cytoscape (v.3.8.2)[103] for network visualization (Supplementary Fig. 9). Module eigengenes were chosen to summarize the gene expression profiles of gene modules. GO enrichment analysis of each gene cluster and gene module was

performed using the topGO (v.2.44) package. We performed a Fisher's exact test and listed the top 30 (soft *k*-means clusters) or top 15 (WGCNA modules) significantly enriched GO terms of the class biological process (Supplementary Tables 27–31, Supplementary Figs. 4–6, 10 and 11). To ease visualization, all 486 non-redundant enriched GO terms from the 33 soft *k*-means clusters from all 3 species were clustered through *k*-means clustering by semantic similarity using the simplifyEnrichment (v.1.2.0) package[104] (Supplementary Figs. 7 and 8). Full network nodes and edges files and the random 30% network subset files are available in the GitHub repository (see Data availability section).

Transcription factor repertoire analysis

We selected a custom set of 36 transcription factor classes from all 9 transcription factor superclasses from the TFClass database[105]. Transcripts in *O. fusiformis*, *C. teleta* and *D. gyrociliatus* were deemed transcription factors and classified into one or more of the 36 classes if they were a match for any of the corresponding PANTHER identifiers (Supplementary Tables 32–33 and Supplementary Fig. 3). Over-representation and under-representation of the different transcription factor classes in the gene expression clusters was tested through pairwise two-tailed Fisher's exact tests, for which we then adjusted the *P* values using Benjamini–Hochberg correction for multiple testing.

Orthogroup overlap analysis

We performed pairwise comparisons between each possible combination of soft *k*-means clusters of all three annelid taxa. The numbers of overlapped orthogroups between either the full clusters or the transcription factors belonging only to each cluster were subjected to upper-tail hypergeometric tests. *P* values were then adjusted using the Benjamini–Hochberg method for multiple testing correction. For the simplified analyses by quadrants, clusters were classed as early/pre-larval (*O. fusiformis*: 1–6; *C. teleta*: 1–5; *D. gyrociliatus*: 1–3) or late/pre-larval (*O. fusiformis*: 8–12; *C. teleta*: 7–12; *D. gyrociliatus*: 5–7), thus rendering 4 different quadrants for each species pairwise comparison: $\text{early}_{\text{species A}}$–$\text{early}_{\text{species B}}$, $\text{early}_{\text{species A}}$–$\text{late}_{\text{species B}}$, $\text{late}_{\text{species A}}$–$\text{early}_{\text{species B}}$ and $\text{late}_{\text{species A}}$–$\text{late}_{\text{species B}}$. Clusters corresponding

to female adult expression in *D. gyrociliatus* (8 and 9) were discarded for comparison purposes. Relative similarity (RS) values for each of the four quadrants were computed as the following ratio:

$${\rm{RS}}=\frac{{\rm{mean}}(-{\log }_{10}{({\rm{adjusted}}P{\rm{value}})}_{{\rm{quadrant}}})}{{\rm{mean}}(-{\log }_{10}{({\rm{adjusted}}P{\rm{value}})}_{{\rm{total}}})}$$

Values above 1 indicate a higher orthogroup overlap than average, whereas values below 1 represent a lower overlap than average. For genes under heterochronic shifts—that is, with distinct temporal expression dynamics—between indirect and direct development, a gene set was constructed with the genes with a single-copy orthologue in both *O. fusiformis* and *C. teleta*, for which expression was shifted from post-larval clusters (*O. fusiformis*: 7–12; *C. teleta*: 8–12) to early clusters 2 and 3 in *D. gyrociliatus* (Fig. 2b, Supplementary Tables 34 and 35 and Supplementary Fig. 12). For the characterization of genes under heterochronic shifts between planktotrophic and lecithotrophic larvae, two gene sets were generated with the genes with $\text{early}_{\textit{O. fusiformis}}$–$\text{late}_{\textit{C. teleta}}$ and $\text{late}_{\textit{O. fusiformis}}$–$\text{early}_{\textit{C. teleta}}$ dynamics, as described above (Supplementary Tables 36–39 and Supplementary Figs. 13 and 14). GO enrichment analysis of both gene sets was performed using the topGO (v.2.44) package. We performed a Fisher's exact test and listed the top 15 significantly enriched GO terms of the class biological process (Supplementary Table 40). BlastKOALA[106] server was used to assign a KEGG orthology number to one-to-one orthologues showing heterochronic sifts and KEGG mapper[107] to analyse the annotations (Supplementary Tables 41 and 42).

Pathway analyses

Human genes involved in the animal autophagy pathway (map04140) were obtained from the KEGG pathway database[108]. *D. melanogaster* and *Saccharomyces cerevisiae* genes involved in the chitin synthesis pathway were fetched from FlyBase[109] and SGD[110], respectively, based on the enzyme nomenclature numbers of the pathway enzymatic activities[111]. Orthology in *O. fusiformis* and *C. teleta* for the autophagy pathway genes was determined from the single-copy orthologue sets to the human genes,

for which one for both species existed (Supplementary Tables 43 and 44). For the chitin synthesis pathway, and owing to the high number of paralogues and expansions and losses of enzymatic activities of the chitin synthesis pathway, orthology was inferred from PANTHER family and subfamily identifiers to the corresponding enzymatic activities (Supplementary Tables 45 and 46). We then used this orthology to reconstruct the chitin synthesis pathway in annelids. Timing across both species and the presence or lack thereof of heterochronic shifts between *O. fusiformis* and *C. teleta* were determined as described above (Supplementary Figs. 15 and 16).

Hox genes orthology assignment

A total of 129 curated Hox sequences were retrieved from various databases (Supplementary Table 47) and aligned with *O. fusiformis* HOX proteins with MAFFT (v.7) in automatic mode. Poorly aligned regions were removed with gBlocks (v.0.91b)[112] to produce the final alignments. Maximum likelihood trees were constructed using RAxML (v.8.2.11.9)[113] with an LG substitution matrix[114] and 1,000 ultrafast bootstraps. All trees were composed in FigTree (v.1.4.4). Alignment files are available in the GitHub repository (see Data availability section).

Whole-mount in situ hybridization and immunohistochemistry

Fragments of *chordin* and Hox genes were isolated as previously described[24] using gene-specific oligonucleotides and a T7 adaptor. Riboprobes were synthesized using a T7 MEGAscript kit (ThermoFisher, AM1334) and stored at a concentration of 50 ng μl^{-1} in hybridization buffer at –20 °C. Whole-mount in situ hybridization in embryonic, larval and juvenile stages were conducted as described elsewhere[24,26]. Antibody staining in larval stages of *O. fusiformis*, *Magelona* spp. and *C. teleta* was carried out as previously described[23,115] using the following antibodies: mouse anti-acetyl-α-tubulin antibody, clone 6-11B-1, 1:800 dilution (Sigma-Aldrich, MABT868, RRID: AB_2819178) and goat anti-mouse IgG (H+L) cross-adsorbed secondary antibody, Alexa Fluor 647, 1:800 dilution (Thermo Fisher Scientific, A-21235, RRID: AB_2535804). Differential interface contrast images of the colorimetric in situ were obtained using a Leica 560

DMRA2 upright microscope equipped with an Infinity5 camera (Lumenera). Fluorescently stained samples were scanned using a Nikon CSU-W1 spinning disk confocal microscope.

ATAC-seq

We performed two replicates of ATAC-seq from samples containing around 50,000 cells at the blastula (about 900 embryos), gastrula (around 500), elongation (about 300), mitraria larva (around 150 larvae) and competent larva (about 40) stages for *O. fusiformis*, and the 64-cells stage (about 500 embryos), gastrula (around 200), stage 4tt larva (about 120 larvae), stage 5 larva (around 90) and stage 8 larva (around 50) for *C. teleta* following the omniATAC protocol[116], but gently homogenizing the samples with a pestle in lysis buffer and incubating them on ice for 3 min. Tagmentation was performed for 30 min at 37 °C with an in-house purified Tn5 enzyme[117]. After DNA clean-up, ATAC-seq libraries were amplified as previously described[116]. Primers used for both PCR and quantitative PCR are listed in Supplementary Tables 57 and 59. Amplified libraries were purified using ClentMag PCR Clean Up beads as indicated by the supplier and quantified and quality checked on a Qubit 4 fluorometer (ThermoFisher) and an Agilent 2200 TapeStation system before pooling at equal molecular weight. Sequencing was performed on an Illumina HiSeq4000 platform in 2 × 75 bp mode at the Oxford Genomics Centre (blastula, elongation and mitraria larva stages, and one replicate of the gastrula sample of *O. fusiformis*, as well as the 64-cells, gastrula and stage 4tt larva stages of *C. teleta*) and on an Illumina NovoSeq6000 in 2 × 150 bp mode at Novogene (one replicate of gastrula and the two replicates of competent larva stages of *O. fusiformis* and the two replicates of stage 5 and stage 8 larva of *C. teleta*).

Chromatin accessibility profiling

We used cutadapt (v.2.5)[118] to remove sequencing adaptors and trim reads from libraries sequenced in 2 × 150 bp mode to 75 bp reads. Quality filtered reads were mapped using NextGenMap (v.0.5.5)[119] in paired-end mode, duplicates were removed using samtools (v.1.9)[120] and mapped reads were shifted using deepTools (v.3.4.3)[121] (Supplementary Tables 58 and 60). Fragment size distribution was estimated from resulting BAM files and transcription start site enrichment analysis was computed using computeMatrix and plotHeatmap commands in deepTools (v.3.4.3). Peak calling was done using MACS2 (v.2.2.7.1)[122,123] (-f BAMPE --min-length 100 --max-gap 75 and -q 0.01). Reproducible peaks were identified by irreproducible discovery rates (values <0.05) (v.2.0.4) at each developmental stage. Peaks from repetitive regions were filtered using BEDtools (v.2.28.0)[124] at each developmental stage. Next we used DiffBind (v.3.0.14)[125] to generate a final consensus peak set of 63,732 peaks in *O. fusiformis* and 46,409 peaks in *C. teleta*, which were normalized using DESeq2 (Supplementary Fig. 17). Peak clustering according to accessibility dynamics was performed as described above for RNA-seq, using the same number of 12 clusters to make both profiling techniques comparable. Principal component analysis and differential accessibility analyses between consecutive developmental stages were also performed as described above. An LFC > 0 and a LFC < 0 indicates whether a peak opens or closes, respectively, given an adjusted P value < 0.05. Stage-specific and constitutive peaks were determined using UpSetR (v.1.4.0)[126], and both the consensus peak set and the stage-specific peak sets were classified by genomic region using HOMER (v.4.11)[127] and further curated. Visualization of peak tracks and gene structures was conducted using pyGenomeTracks (v.2.1)[128] and deepTools (v.3.4.3)[121]. To correlate chromatin accessibility and gene expression, this genomic region annotation was used to assign peaks to their closest gene (63,726 peaks were assigned to 23,025 genes in *O. fusiformis* and 44,368 peaks were assigned to 23,382 genes in *C. teleta*). Pearson correlation coefficient between chromatin accessibility and gene expression was computed individually by peak using two-sided tests (Supplementary Fig. 18). GO enrichment analysis of the

gene sets regulated by peak clusters was performed using the topGO (v.2.44) package. We performed Fisher's exact test and listed the top 30 significantly enriched GO terms of the class biological process (Supplementary Figs. 19 and 20). To ease visualization, all 242 non-redundant enriched GO terms were clustered through *k*-means clustering by semantic similarity using the simplifyEnrichment (v.1.2.0) package[104] (Supplementary Tables 61–71 and Supplementary Figs. 21–23). Coverage files and peak set files are available in the GitHub repository (see Data availability section).

Motif identification, clustering, matching and curation

To identify transcription-factor-binding motifs in chromatin accessible regions in the two species, we first used HOMER[127] (v.4.1) to identify known and de novo motifs in the consensus peak sets, which produced 456 motifs for *O. fusiformis* and 364 motifs for *C. teleta* (Supplementary Tables 72 and 73). Significance of motifs was derived from binomial tests from cumulative binomial distributions. We then used GimmeMotifs (v.0.16.1)[129] with a 90% similarity cut-off to cluster the motifs predicted in *O. fusiformis* and *C. teleta* into 141 consensus motifs, which we matched against four motif databases to assign their putative identity (Gimme vertebrate (5.0)[129], HOMER[127], CIS-BP[130] and a custom JASPAR2022 (ref. [131]) core motifs without plant and fungi motifs; Supplementary Fig. 24). We then used the human non-redundant TF motif database (https://resources.altius.org/~jvierstra/projects/motif-clustering-v2.0beta/) to manually curate the annotation. After removing motifs that probably represented sequence biases, we finally obtained 95 motif archetypes for *O. fusiformis* and 91 for *C. teleta* (Supplementary Table 74), which we then used to perform motif counts in peaks (Supplementary Tables 75 and 76) and motif accessibility estimation (Supplementary Tables 77 and 78) with GimmeMotifs (v.0.16.1)[129]. Data clustering was performed with mfuzz (v.2.52)[101] (Supplementary Figs. 25 and 27). Over-representation and under-representation of counts of the common curated motif archetypes in the peak accessibility soft clusters (see above) was tested through pairwise two-tailed Fisher's exact tests, for which we then adjusted the *P* values using the Bonferroni correction for multiple testing.

Transcription factor footprinting and Hox gene regulatory network exploration

To predict transcription factor binding, as a proxy of activity, we conducted footprinting analysis using TOBIAS (v.0.12.0)[132] during development in the 95 and 91 motif archetypes for *O. fusiformis* and *C. teleta*, respectively (Supplementary Tables 79 and 80). Bound and unbound sites were first estimated by fitting a two-component Gaussian-mixture model, and significance was then tested using a one-tail test from the right-most normal distribution. Transcription factor binding scores (TFBSs) were clustered using mfuzz (v.2.52)[101]. Pearson correlation coefficients of motif accessibility and TFBSs were calculated by stage and by motif separately on the basis of 33 common, curated motif archetypes (Supplementary Figs. 26 and 28–30). To reconstruct potential upstream regulators and downstream effectors of the Hox genes, we first subset ATAC-seq peaks annotated to the Hox genes in the Hox cluster (that is, all except *Post1*) in *O. fusiformis* and *C. teleta* and extracted the bound motifs on those peaks (Supplementary Tables 81 and 82). TFBSs were summed for each motif to obtain global dynamics, and their temporal dynamics were then clustered using mfuzz (v.2.52)[101] (Supplementary Fig. 31). For the downstream genes regulated by Hox, we obtained genes annotated to ATAC-seq peaks with a bound HOX, EVX and CDX motif at the competent stage in *O. fusiformis* and stage 4tt larva in *C. teleta* (Supplementary Tables 83 and 84). One-to-one orthologues were used to identified shared targets and PANTHER identifiers to obtain their functional annotation.

Phylostratigraphy

To evaluate gene expression dynamics by phylostratum and developmental stage in all three annelid lineages, we used the OrthoFinder gene families and their inferred origins. We deemed all genes originating before and with the Cnidarian–Bilaterian ancestor of pre-metazoan and metazoan origin (Supplementary Tables 85–87). We then applied a quantile normalization onto the DESeq2-normalized matrices of gene expression. The 75th percentile of the quantile-normalized gene expression levels was used as the summarizing measure of the gene expression distribution by developmental

stage. Over-representation and under-representation of the different phylostrata in the gene expression clusters were tested through pairwise two-tailed Fisher's exact tests, for which we then adjusted the *P* values using Bonferroni correction for multiple testing. Gene expression dynamics of new genes and genes of pre-metazoan and metazoan origin across selected metazoan lineages (see 'Comparative transcriptomics' section below) were also evaluated as described above (Supplementary Fig. 32).

Comparative transcriptomics

Publicly available RNA-seq developmental time courses for the development of *Amphimedon queenslandica*, *Clytia hemisphaerica*, *N. vectensis*, *S. purpuratus*, *Branchiostoma lanceolatum*, *D. rerio*, *D. melanogaster*, *Caenorhabditis elegans*, *C. gigas*, *D. gyrociliatus*, and two stages of *C. teleta* were downloaded from the SRA using SRA-Toolkit (v.2.11.3) (Supplementary Table 88), cleaned for adaptors and low-quality reads with trimmomatic (v.0.39)[65] and pseudo-aligned to their respective non-redundant genome-based gene repertoires—that is, with a single transcript isoform, the longest, per gene model—using kallisto (v.0.46.2)[99]. We then performed a quantile transformation of TPM values using scikit-learn (v.1.0.2)[133] and calculated the Jensen–Shannon divergence (JSD) value from (1) all single-copy orthologues, (2) the set single-copy transcription factor orthologues and (3) the set of common single-copy orthologues across all lineages, either between all possible one-to-one species comparisons (1) or between all species and *O. fusiformis* (2 and 3), using the philentropy (v.0.5.0) package[134] as follows:

$${{\rm{JSD}}}_{{\rm{raw}}}\left(P\parallel Q\right)=\frac{1}{2}\mathop{\sum }\limits_{i=0}^{n}{p}_{i}\times {\log }_{2}\left(\frac{{p}_{i}}{\frac{1}{2}\left({p}_{i}+{q}_{i}\right)}\right)+\frac{1}{2}\mathop{\sum }\limits_{i=0}^{n}{q}_{i}\times {\log }_{2}\left(\frac{{q}_{i}}{\frac{1}{2}\left({p}_{i}+{q}_{i}\right)}\right)$$

Transcriptomic divergences were calculated on the basis of 250 bootstrap replicates, from which statistically robust mean values and standard deviations were obtained. Raw mean JSD values (JSD_{raw}) were adjusted

(JSD_{adj}) by dividing by the number of single-copy orthologues (1), single-copy transcription factor orthologues (2) or common single-copy orthologues (3) of each comparison (Supplementary Tables 22, 89 and 90) and normalized using the minimum and maximum adjusted JSD values from all one-to-one species comparisons as follows:

$${{\rm{JSD}}}_{{\rm{norm}}}\left(P\parallel Q\right)=\frac{{{\rm{JSD}}}_{{\rm{adj}}}\left(P\parallel Q\right)-\min \,{{\rm{JSD}}}_{{\rm{adj}}}}{\max \,{{\rm{JSD}}}_{{\rm{adj}}}-\min \,{{\rm{JSD}}}_{{\rm{adj}}}}{\rm{;}}{{\rm{JSD}}}_{{\rm{norm}}}\in [0,1]$$

Relative JSD values were obtained equally, using minimum and maximum adjusted JSD values from each one-to-one species comparison instead. Gene-wise JSD (gwJSD) between five key one-to-one larval stages comparisons was computed as follows:

$${\rm{gwJSD}}\left(P\parallel Q\right)=\frac{1}{2}\times {p}_{i}\times {\log }_{2}\left(\frac{{p}_{i}}{\frac{1}{2}\left({p}_{i}+{q}_{i}\right)}\right)+\frac{1}{2}\times {q}_{i}\times {\log }_{2}\left(\frac{{q}_{i}}{\frac{1}{2}\left({p}_{i}+{q}_{i}\right)}\right)$$

Similarity-driving genes—that is, those with very low gwJSD—were subset as those below the threshold defined as 25% of the point of highest probability density of the gwJSD distributions. GO enrichment analysis of the similarity-driving gene sets was performed using the topGO (v.2.44) package. We performed Fisher's exact test and listed the top 30 significantly enriched GO terms of the class biological process (Supplementary Table 91). To ease visualization, all 51 non-redundant enriched GO terms from the 5 gene sets were clustered through *k*-means clustering by semantic similarity using the simplifyEnrichment (v.1.2.0) package[104]. The subsets of similarity-driven transcription factors of each pairwise comparison are listed in Supplementary Table 92. For comparative Hox gene expression dynamics profiling in metazoan lineages, the same non-redundant gene expression matrices were normalized using the DESeq2 (v.1.30.1) package[100] (Supplementary Fig. 33), unless Hox gene

models were missing, in which case they were manually added ad hoc to the non-redundant genome-based gene repertoires (Supplementary Table 94). Hox gene expression profiling in *U. unicinctus* was performed as described for the rest of taxa but using the available reference transcriptome[135] instead (Supplementary Table 48). All gene expression matrices are available in the GitHub repository (see Data availability section).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Accession codes and unique identifiers to previously publicly available datasets we used for this study are listed in Supplementary Table 2 (genome files used in gene family evolution analyses), Supplementary Table 8 (transcriptomes used in the evolutionary analysis of *chordin* in annelids), Supplementary Tables 41 and 43 (gene identifiers used in pathway analyses), Supplementary Table 47 (sequence identifiers used in the orthology assignment of Hox genes), Supplementary Table 48 (RNA-seq datasets used for Hox gene expression profiling in *U. unicinctus*) and Supplementary Table 88 (RNA-seq datasets used for comparative annelid and metazoan transcriptomics and Hox gene expression profiling). Repetitive elements database RepBase can be accessed at https://www.girinst.org/repbase/. Transcription factor public database TFClass can be found at http://tfclass.bioinf.med.uni-goettingen.de/. All sequence data associated with this project are available at the European Nucleotide Archive (project PRJEB38497) and Gene Expression Omnibus (accession numbers GSE184126, GSE202283, GSE192478, GSE210813 and GSE210814). Genome assemblies, TE annotations, genome annotation files used for RNA-seq and ATAC-seq analyses, WGCNA nodes and edges files, alignment files used in orthology assignment and other additional files are publicly available at GitHub (https://github.com/ChemaMD/OweniaGenome).

Code availability

All code used in this study is available at GitHub (https://github.com/ChemaMD/OweniaGenome).

References

1. Hall, B. K. & Wake, M. H. in *The Origin and Evolution of Larval Forms* (eds Hall, B. K. & Wake, M. H.) 1–19 (Academic Press, 1999).

2. Nielsen, C. Animal phylogeny in the light of the trochaea theory. *Biol. J. Linn. Soc.* **25**, 243–299 (2008).

3. Garstang, W. The origin and evolution of larval forms. *Rep. Br. Assoc. Adv. Sci.* **1928**, 77–98 (1928).

4. Jägersten, G. *Evolution of the Metazoan Life Cycle* (Academic Press, 1972).

5. Marlow, H. et al. Larval body patterning and apical organs are conserved in animal evolution. *BMC Biol.* **12**, 7 (2014).

6. Wang, J. et al. Evolutionary transcriptomics of metazoan biphasic life cycle supports a single intercalation origin of metazoan larvae. *Nat. Ecol. Evol.* **4**, 725–736 (2020).

7. Salvini-Plawen, L. V. On the origin and evolution of the lower Metazoa. *J. Zool. Syst. Evol. Res.* **16**, 40–87 (1978).

8. Hazsprunar, G., Salvini-Plawen, L. V. & Rieger, R. M. Larval planktotrophy—a primitive trait in the Bilateria? *Acta Zoologica* **76**, 141–154 (1995).

9. Raff, R. A. Origins of the other metazoan body plans: the evolution of larval forms. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **363**, 1473–1479 (2008).

10. Sly, B. J., Snoke, M. S. & Raff, R. A. Who came first—larvae or adults? Origins of bilaterian metazoan larvae. *Int. J. Dev. Biol.* **47**, 623–632 (2003).

11. Davidson, E. H., Peterson, K. J. & Cameron, R. A. Origin of bilaterian body plans: evolution of developmental regulatory mechanisms. *Science* **270**, 1319–1325 (1995).

12. Wilson, D. P. On the Mitraria larva of *Owenia fusiformis* Delle Chiaje. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **221**, 231–334 (1932).

13. Lacalli, T. C. Protochordate body plan and the evolutionary role of larvae: old controversies resolved? *Can. J. Zool.* **83**, 216–224 (2005).

14. Strathman, R. Multiple origins of feeding head larvae by the Early Cambrian. *Can. J. Zool.* **98**, 761–776 (2020).

15. Gonzalez, P., Uhlinger, K. R. & Lowe, C. J. The adult body plan of indirect developing hemichordates develops by adding a Hox-patterned trunk to an anterior larval territory. *Curr. Biol.* **27**, 87–95 (2017).

16. Gasiorowski, L. & Hejnol, A. Hox gene expression during development of the phoronid *Phoronopsis harmeri*. *EvoDevo* **11**, 2 (2020).

17. Hiebert, L. S. & Maslakova, S. A. Hox genes pattern the anterior–posterior axis of the juvenile but not the larva in a maximally indirect developing invertebrate, *Micrura alaskensis* (Nemertea). *BMC Biol.* **13**, 23 (2015).

18. Rouse, G. W. Trochophore concepts: ciliary bands and the evolution of larvae in spiralian Metazoa. *Biol. J. Linn. Soc.* **66**, 411–464 (2008).

19. Rouse, G. W. Polychaetes have evolved feeding larvae numerous times. *Bull. Mar. Sci.* **67**, 391–409 (2000).

20. Rouse, G. W., Pleijel, F. & Tilic, E. *Annelida* (Oxford Univ. Press, 2022).

21. Smart, T. I. & Von Dassow, G. Unusual development of the mitraria larva in the polychaete *Owenia collaris*. *Biol. Bull.* **217**, 253–268 (2009).

22. Gasiorowski, L. et al. Molecular evidence for a single origin of ultrafiltration-based excretory organs. *Curr. Biol.* **31**, 3629–3638.e2 (2021).

23. Carrillo-Baltodano, A. M., Seudre, O., Guynes, K. & Martin-Duran, J. M. Early embryogenesis and organogenesis in the annelid *Owenia fusiformis*. *EvoDevo* **12**, 5 (2021).

24. Seudre, O., Carrillo-Baltodano, A. M., Liang, Y. & Martín-Durán, J. M. ERK1/2 is an ancestral organising signal in spiral cleavage. *Nat. Commun.* **13**, 2286 (2022).

25. Martin-Duran, J. M. et al. Convergent evolution of bilaterian nerve cords. *Nature* **553**, 45–50 (2018).

26. Martin-Duran, J. M., Passamaneck, Y. J., Martindale, M. Q. & Hejnol, A. The developmental basis for the recurrent evolution of deuterostomy and protostomy. *Nat. Ecol. Evol.* **1**, 5 (2016).

27. Tan, S., Huan, P. & Liu, B. Molluscan dorsal–ventral patterning relying on BMP2/4 and chordin provides insights into spiralian development and evolution. *Mol. Biol. Evol.* **39**, msab322 (2021).

28. Seaver, E. C., Thamm, K. & Hill, S. D. Growth patterns during segmentation in the two polychaete annelids, *Capitella* sp. I and *Hydroides elegans*: comparisons at distinct life history stages. *Evol. Dev.* **7**, 312–326 (2005).

29. Kerbl, A., Fofanova, E. G., Mayorova, T. D., Voronezhskaya, E. E. & Worsaae, K. Comparison of neuromuscular development in two

dinophilid species (Annelida) suggests progenetic origin of *Dinophilus gyrociliatus*. *Front. Zool.* **13**, 49 (2016).

30. Martin-Duran, J. M. et al. Conservative route to genome compaction in a miniature annelid. *Nat. Ecol. Evol.* **5**, 231–242 (2021).

31. Klann, M. & Seaver, E. C. Functional role of *pax6* during eye and nervous system development in the annelid *Capitella teleta*. *Dev. Biol.* **456**, 86–103 (2019).

32. Fröbius, A. C., Matus, D. Q. & Seaver, E. C. Genomic organization and expression demonstrate spatial and temporal *Hox* gene colinearity in the lophotrochozoan *Capitella* sp. I. *PLoS ONE* **3**, e4004 (2008).

33. Pearson, J. C., Lemons, D. & McGinnis, W. Modulating Hox gene functions during animal body patterning. *Nat. Rev. Genet.* **6**, 893–904 (2005).

34. Wei, M. et al. Echiuran Hox genes provide new insights into the correspondence between Hox subcluster organization and collinearity pattern. *Proc. Biol. Sci.* **289**, 20220705 (2022).

35. Seudre, O. et al. The *Fox* gene repertoire in the annelid *Owenia fusiformis* reveals multiple expansions of the foxQ2 class in Spiralia. Preprint at *bioRxiv* https://doi.org/10.1101/2022.03.02.482670 (2022).

36. Boyle, M. J. & Seaver, E. C. Developmental expression of *foxA* and *gata* genes during gut formation in the polychaete annelid, *Capitella* sp. I. *Evol. Dev.* **10**, 89–105 (2008).

37. Wu, L., Ferger, K. E. & Lambert, J. D. Gene expression does not support the developmental hourglass model in three animals with spiralian development. *Mol. Biol. Evol.* **36**, 1373–1383 (2019).

38. Woltereck, R. *Trochophora-Studien I. Histogie der larve und die Entstehung des Annelids bei den Polygordius-Arten der Nordsee* (Verlag von Erwin Nagele, 1902).

39. Irvine, S. Q. & Martindale, M. Q. Expression patterns of anterior Hox genes in the polychaete *Chaetopterus*: correlation with morphological boundaries. *Dev. Biol.* **217**, 333–351 (2000).

40. Peterson, K. J., Irvine, S. Q., Cameron, R. A. & Davidson, E. H. Quantitative assessment of Hox complex expression in the indirect development of the polychaete annelid *Chaetopterus* sp. *Proc. Natl Acad. Sci. USA* **97**, 4487–4492 (2000).

41. Brenneis, G., Bogomolova, E. V., Arango, C. P. & Krapp, F. From egg to "no-body": an overview and revision of developmental pathways in the ancient arthropod lineage Pycnogonida. *Front. Zool.* **14**, 6 (2017).

42. Arenas-Mena, C., Cameron, A. R. & Davidson, E. H. Spatial expression of Hox cluster genes in the ontogeny of a sea urchin. *Development* **127**, 4631–4643 (2000).

43. Arenas-Mena, C., Martinez, P., Cameron, R. A. & Davidson, E. H. Expression of the Hox gene complex in the indirect development of a sea urchin. *Proc. Natl Acad. Sci. USA* **95**, 13062–13067 (1998).

44. Kulakova, M. et al. Hox gene expression in larval development of the polychaetes *Nereis virens* and *Platynereis dumerilii* (Annelida, Lophotrochozoa). *Dev. Genes Evol.* **217**, 39–54 (2007).

45. Aronowicz, J. & Lowe, C. J. Hox gene expression in the hemichordate *Saccoglossus kowalevskii* and the evolution of deuterostome nervous systems. *Integr. Comp. Biol.* **46**, 890–901 (2006).

46. Pascual-Anaya, J. et al. Broken colinearity of the amphioxus Hox cluster. *EvoDevo* **3**, 28 (2012).

47. Carrillo-Baltodano, A. M. & Meyer, N. P. Decoupling brain from nerve cord development in the annelid *Capitella teleta*: insights into the evolution of nervous systems. *Dev. Biol.* **431**, 134–144 (2017).

48. Sur, A., Magie, C. R., Seaver, E. C. & Meyer, N. P. Spatiotemporal regulation of nervous system development in the annelid *Capitella*

teleta. *EvoDevo* **8**, 13 (2017).

49. Nielsen, C. *Animal Evolution. Interrelationships of the Living Phyla* 3rd Edn (Oxford Univ. Press, 2012).

50. Marçais, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of *k*-mers. *Bioinformatics* **27**, 764–770 (2011).

51. Ranallo-Benavidez, T. R., Jaron, K. S. & Schatz, M. C. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat. Commun.* **11**, 1432 (2020).

52. Koren, S. et al. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).

53. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at https://arxiv.org/abs/1303.3997 (2013).

54. Vaser, R., Sovic, I., Nagarajan, N. & Sikic, M. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* **27**, 737–746 (2017).

55. Roach, M. J., Schmidt, S. A. & Borneman, A. R. Purge Haplotigs: allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinformatics* **19**, 460 (2018).

56. Rhie, A., Walenz, B. P., Koren, S. & Phillippy, A. M. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol.* **21**, 245 (2020).

57. Manni, M., Berkeley, M. R., Seppey, M., Simao, F. A. & Zdobnov, E. M. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol. Biol. Evol.* **38**, 4647–4654 (2021).

58. Flynn, J. M. et al. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl Acad. Sci. USA* **117**, 9451–9457 (2020).

59. Buchfink, B., Reuter, K. & Drost, H. G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **18**, 366–368 (2021).

60. Xu, Z. & Wang, H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, W265–W268 (2007).

61. Wong, W. Y. & Simakov, O. RepeatCraft: a meta-pipeline for repetitive element de-fragmentation and annotation. *Bioinformatics* **35**, 1051–1052 (2019).

62. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).

63. Weigert, A. et al. Illuminating the base of the annelid tree using transcriptomics. *Mol. Biol. Evol.* **31**, 1391–1401 (2014).

64. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

65. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

66. Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).

67. Mapleson, D., Venturini, L., Kaithakottil, G. & Swarbreck, D. Efficient and accurate detection of splice junctions from RNA-seq with Portcullis. *GigaScience* **7**, giy131 (2018).

68. Grabherr, M. G. et al. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).

69. Wu, T. D. & Watanabe, C. K. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **21**, 1859–1875 (2005).

70. Venturini, L., Caim, S., Kaithakottil, G. G., Mapleson, D. L. & Swarbreck, D. Leveraging multiple transcriptome assembly methods for improved gene structure annotation. *GigaScience* **7**, giy093 (2018).

71. Stanke, M., Schoffmann, O., Morgenstern, B. & Waack, S. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics* **7**, 62 (2006).

72. Slater, G. S. & Birney, E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* **6**, 31 (2005).

73. Haas, B. J. et al. Improving the *Arabidopsis* genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res.* **31**, 5654–5666 (2003).

74. Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).

75. Bryant, D. M. et al. A tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. *Cell Rep.* **18**, 762–776 (2017).

76. Eddy, S. R. Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).

77. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).

78. Mi, H. et al. PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res.* **49**, D394–D403 (2021).

79. Durand, N. C. et al. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).

80. Dudchenko, O. et al. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* **356**, 92–95 (2017).

81. Shumate, A. & Salzberg, S. L. Liftoff: accurate mapping of gene annotations. *Bioinformatics* **37**, 1639–1643 (2021).

82. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019).

83. Steinegger, M. & Soding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).

84. Huerta-Cepas, J., Serra, F. & Bork, P. ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.* **33**, 1635–1638 (2016).

85. Marletaz, F. et al. Amphioxus functional genomics and the origins of vertebrate gene regulation. *Nature* **564**, 64–70 (2018).

86. Zeng, Q. et al. High-quality reannotation of the king scallop genome reveals no 'gene-rich' feature and evolution of toxin resistance. *Comput. Struct. Biotechnol. J.* **19**, 4954–4960 (2021).

87. Zakas, C., Harry, N. D., Scholl, E. H. & Rockman, M. V. The genome of the poecilogonous annelid *Streblospio benedicti*. *Genome Biol. Evol.* **14**, evac008 (2022).

88. Kwiatkowski, D. et al. The genome sequence of the bootlace worm, *Lineus longissimus* (Gunnerus, 1770). *Wellcome Open Res.* **6**, 272

(2021).

89. Cannon, J. T. et al. Xenacoelomorpha is the sister group to Nephrozoa. *Nature* **530**, 89–93 (2016).

90. Simakov, O. et al. Insights into bilaterian evolution from three spiralian genomes. *Nature* **493**, 526–531 (2013).

91. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

92. Katoh, K. & Frith, M. C. Adding unaligned sequences into an existing alignment using MAFFT and LAST. *Bioinformatics* **28**, 3144–3146 (2012).

93. Whelan, S. & Goldman, N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* **18**, 691–699 (2001).

94. Soubrier, J. et al. The influence of rate heterogeneity among sites on the time dependence of molecular rates. *Mol. Biol. Evol.* **29**, 3345–3358 (2012).

95. Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).

96. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).

97. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).

98. Yang, Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol.*

Evol. **39**, 306–314 (1994).

99. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).

100. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

101. Futschik, M. E. & Carlisle, B. Noise-robust soft clustering of gene expression time-course data. *J. Bioinform. Comput. Biol.* **3**, 965–988 (2005).

102. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).

103. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

104. Gu, Z. & Hübschmann, D. *simplifyEnrichment*: an R/Bioconductor package for clustering and visualizing functional enrichment results. Preprint at *bioRxiv* https://doi.org/10.1101/2020.10.27.312116 (2021).

105. Wingender, E., Schoeps, T., Haubrock, M., Krull, M. & Donitz, J. TFClass: expanding the classification of human transcription factors to their mammalian orthologs. *Nucleic Acids Res.* **46**, D343–D347 (2018).

106. Kanehisa, M., Sato, Y. & Morishima, K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* **428**, 726–731 (2016).

107. Kanehisa, M., Sato, Y. & Kawashima, M. KEGG mapping tools for uncovering hidden features in biological data. *Protein Sci.* **31**, 47–53 (2022).

108. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).

109. Gramates, L. S. et al. FlyBase: a guided tour of highlighted features. *Genetics* **220**, iyac035 (2022).

110. Cherry, J. M. et al. *Saccharomyces* Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res.* **40**, D700–D705 (2012).

111. Merzendorfer, H. The cellular basis of chitin synthesis in fungi and insects: common principles and differences. *Eur. J. Cell Biol.* **90**, 759–769 (2011).

112. Talavera, G. & Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**, 564–577 (2007).

113. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).

114. Le, S. Q. & Gascuel, O. An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25**, 1307–1320 (2008).

115. Meyer, N. P., Carrillo-Baltodano, A., Moore, R. E. & Seaver, E. C. Nervous system development in lecithotrophic larval and juvenile stages of the annelid *Capitella teleta*. *Front. Zool.* **12**, 15 (2015).

116. Corces, M. R. et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **14**, 959–962 (2017).

117. Picelli, S. et al. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* **24**, 2033–2040 (2014).

118. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* https://doi.org/10.14806/ej.17.1.200

(2011).

119. Sedlazeck, F. J., Rescheneder, P. & von Haeseler, A. NextGenMap: fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics* **29**, 2790–2791 (2013).

120. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

121. Ramirez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).

122. Gaspar, J. M. Improved peak-calling with MACS2. Preprint at *bioRxiv* https://doi.org/10.1101/496521 (2018).

123. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

124. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

125. Ross-Innes, C. S. et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).

126. Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938–2940 (2017).

127. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576–589 (2010).

128. Lopez-Delisle, L. et al. pyGenomeTracks: reproducible plots for multivariate genomic datasets. *Bioinformatics* **37**, 422–423 (2020).

129. van Heeringen, S. J. & Veenstra, G. J. GimmeMotifs: a de novo motif prediction pipeline for ChIP-sequencing experiments. *Bioinformatics* **27**, 270–271 (2011).

130. Weirauch, M. T. et al. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **158**, 1431–1443 (2014).

131. Castro-Mondragon, J. A. et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **50**, D165–D173 (2022).

132. Bentsen, M. et al. ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. *Nat. Commun.* **11**, 4267 (2020).

133. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

134. Drost, H. G. Philentropy: information theory and distance quantification with R. *J. Open Source Softw.* **3**, 765 (2018).

135. Park, C. et al. The developmental transcriptome atlas of the spoon worm *Urechis unicinctus* (Echiurida: Annelida). *GigaScience* **7**, giy007 (2018).

Acknowledgements

We thank A. de Mendoza, D. Gavriouchkina and S. Rossiter for their support and valuable comments on the manuscript; staff at Station Biologique de Roscoff for their help with collection and animal supplies; staff at the Oxford Genomics Centre at the Wellcome Centre for Human Genetics (funded by Wellcome Trust grant reference 203141/Z/16/Z) for the generation and initial processing of RNA-seq and ATAC-seq sequencing data; M. J. Boyle for providing the *chordin* sequence for *T. lageniformis*; J. Deane for initial help with Hox gene characterization in *O. fusiformis*; and core technical staff at the Department of Biology at Queen Mary University of London for their constant support. This work used computing resources from Queen Mary University of London's Apocrita HPC facilities. This work was funded by the Horizon 2020 Framework Programme to J.M.M.-D. (European Research Council Starting Grant action number 801669) and A.H. (European Research Council Consolidator Grant action number

648861), a Royal Society University Research Fellowship (URF\R1\191161) and a Japan Society for the Promotion of Science Kakenhi grant (JP 19K06620) to F.M., and a Biotechnology and Biological Sciences Research Council LIDo iCASE PhD studentship (BB/T008709/1) to J.M.M.-D. and B.E.D.

Author information

Author notes

1. These authors contributed equally: Francisco M. Martín-Zamora, Yan Liang

Authors and Affiliations

1. School of Biological and Behavioural Sciences, Queen Mary University of London, London, UK

   Francisco M. Martín-Zamora, Yan Liang, Kero Guynes, Allan M. Carrillo-Baltodano, Billie E. Davies, Rory D. Donnellan, Giacomo Moggioli, Océane Seudre, Martin Tran & José M. Martín-Durán

2. Genomics and Regulatory Systems Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan

   Yongkai Tan & Nicholas M. Luscombe

3. Department of Infectious Disease, Imperial College London, London, UK

   Martin Tran

4. Department of Natural Sciences, Amgueddfa Cymru–Museum Wales, Cardiff, UK

   Kate Mortimer

5. Department of Biological Sciences, University of Bergen, Bergen, Norway

Andreas Hejnol

6. Institute of Zoology and Evolutionary Research, Faculty of Biological Sciences, Friedrich Schiller University Jena, Jena, Germany

Andreas Hejnol

7. Department of Genetics, Evolution and Environment, University College London, London, UK

Ferdinand Marlétaz

Contributions

J.M.M.-D., F.M., Y.L. and F.M.M.-Z. conceived and designed the study. Y.L. collected RNA-seq samples for *O. fusiformis* and *C. teleta*, performed ATAC-seq experiments and contributed to all data analyses. F.M.M.-Z. performed *chordin* orthology studies and contributed to all data analyses. K.G. conducted in situ hybridization analyses of Hox genes. A.M.C.-B. collected RNA-seq samples for *C. teleta*, performed immunostainings on larvae and gene expression analyses of *chordin*. B.E.D. and R.D.D. contributed to computational analyses. Y.T. generated OMNI-C libraries. G.M. performed repeat annotations and analyses. O.S. identified and performed in silico analyses of Hox genes. M.T. performed genomic extractions and optical mapping. K.M. collected *Magelona* spp. A.H. and N.M.L. contributed to sequencing efforts. F.M. and J.M.M.-D. assembled and annotated the genome and contributed to data analyses. Y.L., F.M.M.-Z. and J.M.M.-D. drafted the manuscript, and all authors critically read and commented on the manuscript.

Corresponding authors

Correspondence to Ferdinand Marlétaz or José M. Martín-Durán.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Bernard Degnan, Kevin Kocot, Greg Rouse and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 The genome of *Owenia fusiformis* is conservatively evolving.

a, Differential interface contrast (DIC) images and z-stack confocal laser scanning views of a *C. teleta* trochophore larva stained for DAPI and acetylated α-tubulin. **b,c**, Pie charts of the transposable element content and Kimura substitution plots of transposable element divergence for *O. fusiformis* and other selected annelid species belonging to different annelid clades as depicted in **c**. Unlike *H. robusta* and *L. luymesi*, which show bursts of transposable elements, *O. fusiformis* shows more steady rates of expansion. **d**, Gene family evolution analysis across 22 metazoan lineages under a consensus tree topology. Gains are shown in green, losses in violet. Gene family losses in *O. fusiformis* are like those of slow-evolving lineages. **e**, Principal component analysis from Fig. 1b, showing the full set of

species. **f,g**, *O. fusiformis* has the lowest number of gene losses of all sampled annelids (**e**), and the least gene expansions (**f**) after the extremely compact genome of *D. gyrociliatus*. **h**, Macrosynteny analysis between *O. fusiformis*, and from top to bottom, the cephalochordate *Branchiostoma floridae*, the bivalve *Pecten maximus*, and the annelid *Streblospio benedicti*. *Owenia fusiformis* retains ancestral linkage groups but also exhibits annelid- and species-specific chromosomal arrangements. However, the karyotype of *O. fusiformis* is more conserved than that of the annelid *S. benedicti*. **i**, Macrosynteny analysis between the bivalve *P. maximus* and the nemertean worm *L. longissimus*. *Lineus longissimus* exhibits conserved ancestral bilaterian linkage groups, including three potential lophotrochozoan-specific chromosomal rearrangements (H+Q, J2+L and K+O2), plus a nemertean-specific fusion (G+C1). Scale bar in **a**, 50 µm.

Extended Data Fig. 2 *chordin* was lost multiple times in annelids.

a, Domain organisation of Chordin (CHRD) and Chordin-like (CHRDL1/2) proteins, as inferred from human orthologs. **b**, Public AlphaFold protein structure prediction for human Chordin (UniProt: Q9H2X0) and Chordin-like 1 (UniProt: Q9BU40) revealed a previously unknown and uncharacterised domain in CHRDL1 and CHRDL2 (also depicted in **a**). **c**,**d**, Orthology assignment of *chordin* annelid candidates. From the multiple sequence alignment, candidate annelid sequences with a 10-residue or longer fragment aligned against either the CHRD (**c**; i.e., bona fide *chordin* genes) or the vWFC domains (**d**; i.e., putative *chordin* genes) were kept for further analysis. CHRDL cluster is shaded in blue; CHRD cluster, in red. Bootstrap support values (top) and posterior probabilities (bottom) are shown at both key nodes. Sequences in red and blue are curated CHRD and CHRDL sequences, respectively. **e,f**, Summary phylogenetic trees of presence or absence of *chordin* (red) or putative *chordin* (light brown) across Annelida. **g**, RNA-seq expression levels of *chordin* in *O. fusiformis*, which peaks at the blastula and gastrula stages, after the specification and inductive activity of the embryonic organiser. Curve is a locally estimated scatterplot smoothing, coloured shaded area represents standard error of the mean. **h**, Whole mount *in situ* hybridisation of *chordin* at the blastula (5 h post fertilisation, hpf), gastrula (9 hpf), and mitraria larva (27 hpf) stages of

O. fusiformis. Asterisks mark the animal/anterior pole. gp: gastral plate; bp: blastopore, mo: mouth. Representative results of three independent analyses. Scale bar in **h**, 50 µm.

Extended Data Fig. 3 Gene expression dynamics during annelid life cycles.

a,b, Heatmaps of average pairwise transcriptomic Jensen–Shannon Divergence (JSD) between *O. fusiformis* and *C. teleta* (**a**), and between *D. gyrociliatus* and either *O. fusiformis* (**b**, left) or *C. teleta* (**b**, right). Average relative JSD of the *C. teleta* or *O. fusiformis* stages of minimal divergence to each corresponding stage is shown on top. Confidence intervals represent standard deviation from 250 bootstrap resamplings of the ortholog sets. **c–e**, Soft *k*-means clustered heatmap of all transcripts whose expression was not null in at least one developmental stage into an optimal number of 12 clusters (*O. fusiformis*, **c**; and *C. teleta*, **d**) and 9 clusters (*D. gyrociliatus*, **e**). Soft clustering considerably increased temporal resolution for the RNA-seq time course of *D. gyrociliatus*. On the right of each heatmap, gene-wise expression dynamics (grey lines) and locally estimated scatterplot smoothing (coloured lines) for each cluster. Coloured shaded areas represent standard error of the mean. **f**, Enrichment analysis of biological process gene ontology (GO) terms for RNA-seq clusters. Each line represents a single GO term, for which the $-\log_{10}(p\text{–value})$ for each RNA-seq cluster is shown in a colour-coded scale. GO terms were clustered into 15 distinct clusters based on semantic similarity (see Supplementary Figs. 7, 8). Clusters are shown on the bottom of the heatmaps. For the full list of GO terms and clusters, see Supplementary Figs. 4–6. *P*-values were derived from upper-tail Fisher's exact tests.

Extended Data Fig. 4 Heterochronic shifts in gene regulatory programmes between annelid life cycles.

a,b, Similarity heatmaps showcasing the orthogroup overlap between the clusters of co-regulated genes (see Extended Data Fig. 3c–e), between the three annelids. *P*-values were derived from upper-tail hypergeometric tests and Benjamini-Hochberg-adjusted. **c**, Explanation of the orthogroup

overlap analysis by quadrants. Clusters were classed as “early” (before dotted lines) or “late” (after dotted lines). Clusters of the female adult of *D. gyrociliatus* were disregarded. **d,e**, Heatmaps of relative similarity by quadrants of the orthogroup overlap analyses of the whole genomes (**d**) and transcription factors only (**e**). Colour scale in **d** and **e** is the same as in **c**. **f**, KEGGbrite characterisation of the gene sets under heterochronic shifts (surrounded by dotted black lines in Fig. 2e) between *O. fusiformis* and *C. teleta*. **g**, Bar plots depicting *p*-values of top biological process GO terms of genes shifted from late expression in *O. fusiformis* to early expression in *C. teleta*. *P*-values were derived from upper-tail Fisher’s exact tests. Full list is available in Supplementary Fig. 13. **h**, Enrichment analysis of the number of transcription factors per class in clusters of co-transcribed genes of *O. fusiformis* (left), *C. teleta* (centre) and *D. gyrociliatus* (right). For each cluster and class combination, the Bonferroni-adjusted *p*–value from the two-sided Fisher’s exact test is shown. Cells in red represent overrepresented classes (odds ratio, OR > 1; adjusted *p*–value < 0.05); cells in blue, underrepresented classes (OR < 1, adjusted *p*–value < 0.05). Dotted lines highlight clusters of maximal enrichment of the homeodomain class. n.s.: not significant.

Extended Data Fig. 5 The *Hox* gene complement and expression in *O. fusiformis*.

a, Orthology assignment of *O. fusiformis Hox* genes through maximum likelihood phylogenetic inference. Bootstrap support values are shown for major gene groups. Of: *O. fusiformis*. **b**, Volcano plot of the mitraria to competent larva transition, highlighting the marked upregulation of *Hox* genes. LFC: $\log_2$(fold-change). *P*-values were derived from the described DESeq2 pipeline and Benjamini-Hochberg-adjusted. **c**, Chromosomal location of the *Hox* cluster and *Post1* gene in *O. fusiformis* (top) and schematic comparison of *Hox* cluster organisation in annelids and a mollusc (bottom). Arrows denote direction of transcription. **d**, Schematic representation to scale of the genomic loci and intron–exon composition of *Hox* genes in *O. fusiformis*. **e**, Heatmaps of *Hox* gene expression during the development of *C. teleta*, *O. fusiformis* and the echiuran annelid *Urechis unicinctus*. In the two annelid species with planktotrophic larvae, *Hox* genes only become expressed at the larval stage (dotted vertical line), and not

during embryogenesis, as observed in *C. teleta*. **f**, Whole mount *in situ* hybridisation of *Hox* genes in the gastrula (lateral views) and in the mitraria larva, pre-competent larva, and juvenile stages of *O. fusiformis* (ventral views). The area encircled by a dotted white line at the pre-competent stage highlights a region of probe trapping from ingested food content. bp: blastopore; mo: mouth. Representative results of three independent analyses. **g,h** Schematic representations of the expression of *Hox* genes in the trunk rudiment of the competent larva (**g**) and juvenile trunk (**h**). A: anterior; P: posterior. Drawings are not to scale, and schematic expression domains are approximate. Scale bars in **f**, 50 μm in gastrulae and larvae, and 100 μm in juvenile.

Extended Data Fig. 6 Transcriptomic dynamics of anteroposterior genes.

a, Schematic drawing of the adult body regions used to define anterior and posterior and trunk genes. **b**, Correlation matrix of RNA-seq experiments from all nine adult tissues, calculated from a variance stabilising-transformed matrix of the normalised DESeq2 matrix. **c**, Venn diagram showing the number of tissue-specific and shared expressed genes (TPM > 2). Gene sets highlighted with red text were defined as adult anterior, and adult posterior and trunk genes. **d**, Phylostratigraphic classification of adult anterior, and adult posterior and trunk genes, compared to the whole genome and a random subset of 1,000 genes. **e,f**, Expression dynamics of each phylostratum by developmental stage in the adult anterior (**e**), and adult posterior and trunk gene sets (**f**), calculated from the 75 % percentile of a quantile-normalised matrix of gene expression levels. Adult anterior genes of most phylostrata peak at the blastula, while the maximum expression of adult trunk/posterior genes of most phylostrata peak at post-larval stages. **g–l**, Average expression dynamics of *in situ* hybridisation-validated anterior, trunk, and posterior markers throughout *O. fusiformis* (**g,h**), *C. teleta* (**i,j**), and *D. gyrociliatus* (**k,l**) development. For boxplots in **g**,**i**, and **k**, centre lines, median; box, interquartile range (IQR); whiskers, first or third quartile ± 1.5 × IQR. Lower whiskers are sometimes not apparent due to the distribution skewness towards zero. Curves in **h**,**j**, and **l** are locally estimated scatterplot smoothings. Coloured shaded areas represent standard error of the mean. *n* = 23, 8, and 17 anterior markers, 10

and 3 posterior markers, and 15, 10, and 8 trunk markers, for *O. fusiformis*, *C. teleta*, and *D. gyrociliatus*, respectively. Key stages where expression of trunk markers is incipient are shown for both *O. fusiformis* and *C. teleta*.

Extended Data Fig. 7 Chromatin dynamics during annelid development.

a, Genomic feature annotation of the consensus ATAC-seq peaks. **b**, Stacked bar plots showing the proportion of called peaks per developmental stage classified by genomic feature. **c**, Heatmap of normalised peak accessibility of the soft clustered consensus ATAC-seq peak sets. **d**, Self-correlation matrices of normalised motif accessibility and transcription factor binding score, revealing distinct chromatin regulatory dynamics throughout development. **e**, Correlation matrices of normalised motif accessibility to transcription factor binding score during annelid development. **f**, Correlation matrix of normalised motif accessibility between both species. **d–f** further validate the non-triviality of the results obtained in Fig. 3c. Pearson correlation coefficients in **d–f** were derived from two-tailed tests. **g**, Heatmap of normalised motif accessibility and transcription factor binding dynamics for each of the common annotated annelid motif archetypes during *O. fusiformis* and *C. teleta* development. Colour scale denotes transcription factor binding score dynamics, bubble size represents motif accessibility dynamics, both in a z-score scale. Motif archetypes highlighted in red are representative examples of the heterochronic shifts shown in bulk in Fig. 3c. **h**, Enrichment analysis of the number of occurrences of the common annotated annelid motif archetypes in the peak clusters inferred through soft *k*-means clustering and shown in **c**, for *O. fusiformis* (top) and *C. teleta* (bottom). For each cluster and motif combination, the Bonferroni-adjusted *p*–value of the two-tailed Fisher's exact test is shown. Red cells represent significantly overrepresented lineages (odds ratio, OR > 1, adjusted *p*–value < 0.05). Blue cells denote significantly underrepresented lineages (OR < 1, adjusted *p*–value < 0.05). **i**, Most abundant bound motifs in peaks of the *Hox* clusters. **j**, Downstream regulated genes by transcription factors bound to the HOX/CDX/EVX motif archetype.

Extended Data Fig. 8 Phylostratigraphy analyses in annelid life cycles.

a,**b** Expression dynamics (**a**) and expression contribution (**b**) of each phylostratum by developmental stage in all three annelids, calculated from the 75% percentile of a quantile-normalised matrix of gene expression levels. Older genes are expressed at the highest levels across annelid development. **c–e**, Boxplots of quantile-normalised expression levels of genes classified by phylostratum across *O. fusiformis* (**c**), *C. teleta* (**d**), and *D. gyrociliatus* (**e**) development. A random subset of 2,000 genes is shown as a negative control. *n* denotes number of genes per phylostratum. **f–h**, Enrichment analysis of the number of genes per phylostratum in clusters of co-transcribed genes as inferred through soft *k*-means clustering and shown in Extended Data Fig. 3c–e, for *O. fusiformis* (**f**), *C. teleta* (**g**), and *D. gyrociliatus* (**h**). For each cluster and phylostratum combination, the Bonferroni-adjusted *p*–value of the two-tailed Fisher's exact test is shown. Upper tables include significantly overrepresented lineages (odds ratio, OR > 1, adjusted *p*–value < 0.05). Lower tables include significantly underrepresented lineages (OR < 1, adjusted *p*–value < 0.05). Shaded grey areas indicate clusters of genes with peak expression at the mitraria larva, for *O. fusiformis*; and stage 4tt through stage 7 larval stages, for *C. teleta*.

Extended Data Fig. 9 Bilaterian planktotrophic larvae and cnidarian larvae share maximal transcriptional similarity.

a, Heatmaps of normalised transcriptomic Jensen–Shannon divergence (JSD) from pairwise comparisons of all single copy one-to-one orthologs (left), the set of common orthologs to all species (centre), and all single copy one-to-one transcription factor orthologs (right), between *O. fusiformis* and ten other metazoan lineages with different life cycles. Larval stages are highlighted in green. **b**, Average relative JSD for the datasets shown in **a**, from stages of minimal JSD to each *O. fusiformis* stage. Confidence intervals represent the standard deviation from 250 bootstrap resamplings of the ortholog sets. **c–e**, Stages of minimal JSD to each *O. fusiformis* stage, calculated from the one-to-one ortholog set (**c**), the common ortholog set (**d**), and the one-to-one transcription factor ortholog

set (**e**). Larval stages are highlighted in green. **f**, Violin plots of the gene-wise Jensen Shannon divergence (gwJSD) distributions for the pairwise comparisons of the one-to-one ortholog sets between the mitraria larva of *O. fusiformis* and the stages of minimal transcriptomic divergence as in **c**. for *C. gigas* (*n* = 6,737 single copy orthologs), *C. hemisphaerica* (*n* = 4,691), *C. teleta* (*n* = 7,651), *N. vectensis* (*n* = 5,254), and *S. purpuratus* (*n* = 5,015). Boxes represent mean estimate ± standard deviation. Dotted lines mark the point of highest probability density. Genes below ¼ of this point were subset as similarity-driving genes. **g**, Biological process GO terms enrichment of the five similarity-driving gene sets. GO terms were clustered by semantic similarity into 4 clusters. Each row represents a single GO term, for which the $-\log_{10}(p\text{–value})$ for each gene set is shown in a colour-coded scale.

Extended Data Fig. 10 Comparative transcriptomic analysis of metazoan life cycles.

a, Matrix of heatmaps of normalised transcriptomic Jensen–Shannon divergence (JSD) from pairwise comparisons of all single copy one-to-one orthologs between all eleven metazoan lineages. From top to bottom and left to right: the annelids *O. fusiformis* and *C. teleta*, the bivalve *C. gigas*, the nematode *C. elegans*, the insect *D. melanogaster*, the vertebrate *D. rerio*, the cephalochordate *B. lanceolatum*, the sea urchin *S. purpuratus*, the cnidarians *N. vectensis* and *C. hemisphaerica*, and the poriferan *A. queenslandica*. **b**, Proposed evolutionary scenario for larval and life cycle evolution in Annelida. Post-embryonic trunk patterning is likely an ancestral condition with the convergent pre-displacement of trunk differentiation to embryogenesis concurring with the evolution of indirect development with feeding larva and direct development. Drawings are not to scale. **c**, Expression dynamics of *Hox* genes across the developmental RNA-seq time courses of all eleven species from **a** and the echiuran annelid *U. unicinctus*. Heatmaps were vertically aligned at the blastula, gastrula, and juvenile stages for all species. Lophotrochozoan lineages with trochophore larvae were also vertically aligned at the trochophore stage. Dotted lines encompass the larval stages of species with ciliated larvae. See Extended Data Fig. 5e and Supplementary Fige. 33 for the fully labelled and non-deformed heatmaps. **d**, Alternative evolutionary scenarios for the

deployment of *Hox* genes (as proxy for trunk patterning and assuming the staggered expression along the directive axis of cnidarians and anteroposterior axis of bilaterians is homologous, which does not necessarily imply homology of the two axes). Given our current understanding of *Hox* gene deployment in cnidarian and bilaterian taxa, a late post-embryonic *Hox* patterning ancestral to Bilateria and Cnidaria, as seen in extant lineages with maximal indirect development, is a more parsimonious scenario (on the right).

Supplementary information

Supplementary Information

This file contains 33 Supplementary Figs. and legends for 94 Supplementary Tables.

Reporting Summary

Peer Review File

Supplementary Tables

Supplementary Tables 1–94.

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain

permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Martín-Zamora, F.M., Liang, Y., Guynes, K. *et al.* Annelid functional genomics reveal the origins of bilaterian life cycles. *Nature* **615**, 105–110 (2023). https://doi.org/10.1038/s41586-022-05636-7

- Received: 10 February 2022
- Accepted: 07 December 2022
- Published: 25 January 2023
- Issue Date: 02 March 2023
- DOI: https://doi.org/10.1038/s41586-022-05636-7

Share this article

Anyone you share the following link with will be able to read this content:

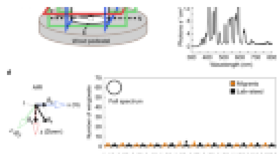
Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

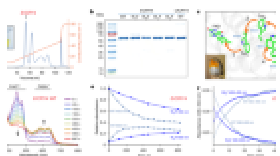
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-022-05636-7



Cryptochrome 1 mediates light-dependent inclination magnetosensing in monarch butterflies

03 February 2021

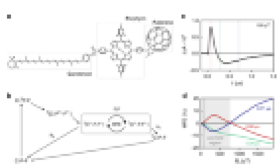
Guijun Wan, Ashley N. Hayden, … Christine Merlin



Magnetic sensitivity of cryptochrome 4 from a migratory songbird

23 June 2021

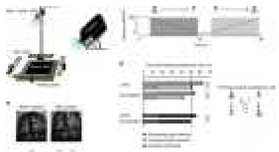
Jingjing Xu, Lauren E. Jarocha, … P. J. Hore



Chemical compass behaviour at microtesla magnetic fields strengthens the radical pair hypothesis of avian magnetoreception

16 August 2019

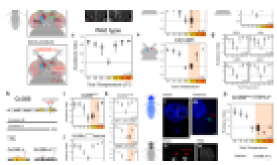
Christian Kerpal, Sabine Richert, … Christiane R. Timmel



[Dedicated photoreceptor pathways in Drosophila larvae mediate navigation by processing either spatial or temporal cues]

28 March 2018

Tim-Henning Humberg, Pascal Bruegger, … Simon G. Sprecher



[Robustness and plasticity in Drosophila heat avoidance]

06 April 2021

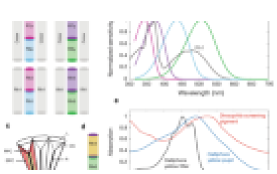
José Miguel Simões, Joshua I. Levy, … Marco Gallio



[Magnetic compass of garden warblers is not affected by oscillating magnetic fields applied to their eyes]

26 February 2020

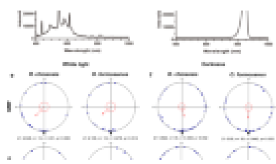
Julia Bojarinova, Kirill Kavokin, … Nikita Chernetsov



[The spectral sensitivity of Drosophila photoreceptors]

26 October 2020

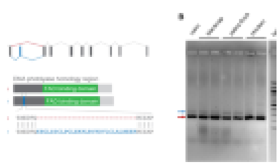
Camilla R. Sharkey, Jorge Blanco, … Trevor J. Wardill



[In addition to cryptochrome 2, magnetic particles with olfactory co-receptor are important for magnetic orientation in termites]

23 September 2021

Yongyong Gao, Ping Wen, … Qiuying Huang



[A novel isoform of cryptochrome 4 (Cry4b) is expressed in the retina of a night-migratory songbird]

25 September 2020

Angelika Einwich, Karin Dedek, … Henrik Mouritsen

- Article
- Open Access
- Published: 22 February 2023

Essential elements of radical pair magnetosensitivity in *Drosophila*

- Adam A. Bradlaugh,
- Giorgio Fedele,
- Anna L. Munro,
- Celia Napier Hansen,
- John M. Hares,
- Sanjai Patel,
- Charalambos P. Kyriacou,
- Alex R. Jones,
- Ezio Rosato &
- …
- Richard A. Baines

Show authors

Nature **volume 615**, pages 111–116 (2023)

- 7124 Accesses
- 354 Altmetric
- Metrics details

Abstract

Many animals use Earth's magnetic field (also known as the geomagnetic field) for navigation[1]. The favoured mechanism for magnetosensitivity involves a blue-light-activated electron-transfer reaction between flavin adenine dinucleotide (FAD) and a chain of tryptophan residues within the photoreceptor protein CRYPTOCHROME (CRY). The spin-state of the resultant radical pair, and therefore the concentration of CRY in its active state, is influenced by the geomagnetic field[2]. However, the canonical CRY-centric radical-pair mechanism does not explain many physiological and behavioural observations[2,3,4,5,6,7,8]. Here, using electrophysiology and behavioural analyses, we assay magnetic-field responses at the single-neuron and organismal levels. We show that the 52 C-terminal amino acid residues of *Drosophila melanogaster* CRY, lacking the canonical FAD-binding domain and tryptophan chain, are sufficient to facilitate

magnetoreception. We also show that increasing intracellular FAD potentiates both blue-light-induced and magnetic-field-dependent effects on the activity mediated by the C terminus. High levels of FAD alone are sufficient to cause blue-light neuronal sensitivity and, notably, the potentiation of this response in the co-presence of a magnetic field. These results reveal the essential components of a primary magnetoreceptor in flies, providing strong evidence that non-canonical (that is, non-CRY-dependent) radical pairs can elicit magnetic-field responses in cells.

Main

The ability of species to navigate considerable distances has long intrigued the biological community[1]. One of several environmental cues to support these migrations is the geomagnetic field (geoMF). Moreover, several other behaviours respond reliably to magnetic fields (MFs), at least under laboratory conditions, showing that the ability to sense and react to MFs is not limited to migrating animals[9]. However, the identity of the primary magnetoreceptors, the mechanisms that underlies their reported light dependence and how the magnetic signal is transduced remain unclear[10,11]. A favoured model posits a light-induced electron-transfer reaction whereby radical pairs (RPs) are formed, the spin-states of which are sensitive to MFs as small as the geoMF (around 50 μT)[2]. This so-called RP mechanism (RPM) canonically requires the flavoprotein CRY, which is best known for its role as a circadian blue light (BL) photoreceptor in flies and as a light-insensitive transcriptional regulator in the circadian clock of mammals[2,10].

Absorption of BL by CRY-bound FAD initiates an electron-transfer cascade along a conserved chain of tryptophan (Trp) residues[2,12,13,14]. In *Drosophila* this forms a spin-correlated RP comprising the photoreduced FAD (FAD^{o-}) and the terminal oxidized Trp ($TrpH^{o+}$)[15]. The spin-state of the RP is initially polarized as a singlet (S, anti-parallel spins), which then rapidly oscillates between S and the triplet spin states (T, parallel spins). Transiently (that is, before the system relaxes to equilibrium), this interconversion can be sensitive to MFs, which in turn can lead to downstream modifications in the biological activity of *Drosophila melanogaster* CRY (*Dm*CRY) through conformational change[2]. In its activated state, the *Dm*CRY C-terminal tail of

around 20 residues (CTT) becomes exposed, enabling interactions with signalling partners, including PDZ-domain-containing proteins[16,17,18,19,20,21,22].

Although there is ample evidence consistent with CRY being both necessary and sufficient for light-dependent magnetosensitivity, there are a number of studies that support exceptions to this mechanism[2,3,4,5,6,7,8]. In one of the most notable, a circadian behavioural assay in *Drosophila* was used to show that *Dm*CRY-dependent light and magnetosensitivity could be rescued in Dm*cry*-null adult flies through expression of the 52 C-terminal (CT) residues of *Dm*CRY fused to GFP (GFP–CT) for stability[3]. Furthermore, *Dm*CRYΔ, resulting from the deletion of the CTT of *Dm*CRY, appeared largely insensitive to an MF, although BL sensitivity was maintained[3,4].

The *Drosophila* CRY CT lacks both the FAD-binding pocket and the chain of four Trp residues (Trp394, Trp342, Trp397 and Trp420) that are presumed to be necessary for the canonical RPM[2,23,24,25]. Moreover, mutating these Trp residues, including W420F and W342F, at best attenuates, but does not abolish, the magnetic functionality of *Dm*CRY[3,4,26,27]. These results are inconsistent with the current understanding of the RPM and question the identity of the magnetically sensitive RP in the receptor. Proposed alternatives to a RP between FAD^{o-} and $TrpH^{o+}$ include the formation of an RP between FAD^{o-}/$FADH^{o}$ and O_2^{o-} or another (unknown, Z^{o}) radical. It is a matter of some contention whether these unconventional RPs contribute to magnetoreception or even represent a primary sensor[11,28,29,30].

Here we report the expression of a new transgene encoding *Dm*CRY-CT fused to luciferase (hereafter, Luc–CT), which lacks the canonical FAD-binding pocket and Trp chain and is therefore unable to support light-induced intramolecular electron transfer. Nevertheless, Luc–CT was sufficient to generate changes in BL- and MF-dependent phenotypes in a whole-organism circadian behavioural assay and in the electrophysiological activity of a model neuron, the larval aCC motorneuron. We show that the MF responsiveness of Luc–CT is potentiated by increasing the intracellular concentration of free FAD, to the point at which high levels of this flavin alone can, in the absence of Luc–CT, support an MF response. Finally, we confirm by mutational analysis that the integrity of the CTT of *Dm*CRY

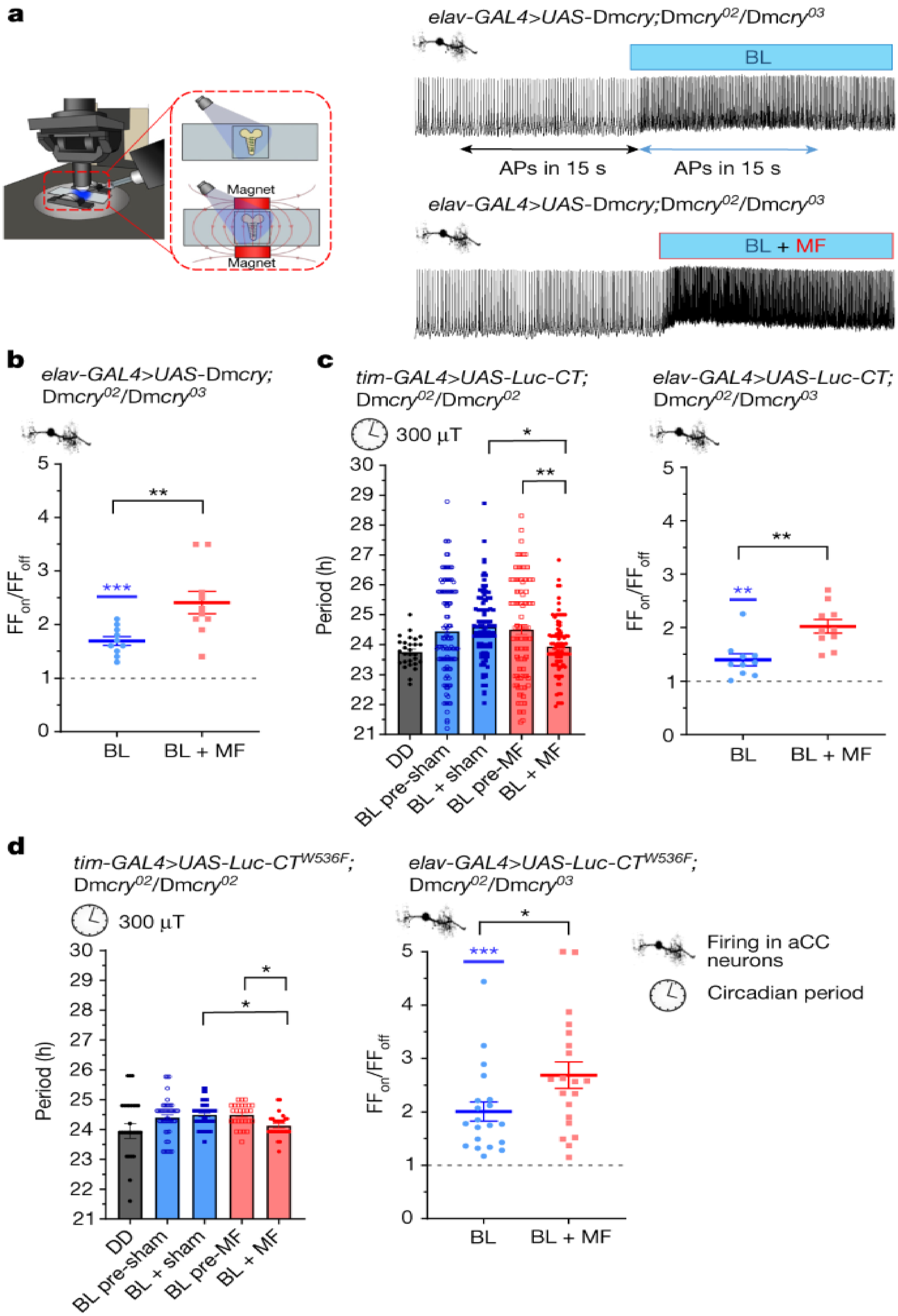
correlates with its ability to facilitate sensitivity to an MF. Overall, our results suggest that 'sensing' and 'transducing' MFs are separate properties that do not need to be carried out by the same molecule.

*Dm*CRY-CT supports magnetosensitivity

To validate our electrophysiological assay, we expressed a full-length Dm*cry* transgene in the aCC motorneuron; this supported a BL-induced increase in action potential firing by 1.7-fold and by 2.4-fold, in the co-presence of an MF[4] (BL + MF, 100 mT static, $P = 0.005$; Fig. 1a,b and Extended Data Fig. 1a). Expression of *Dm*CRY-CT (fused to GFP) in a Dm*cry-null* background supports an MF-induced shortening of the circadian period[3]. To eliminate the possibility that GFP might, like *Dm*CRY, support intramolecular light-induced electron-transfer, we fused *Dm*CRY-CT to luciferase and maintained the flies in the absence of luciferin substrate. BL lengthened the free-running period of *tim-GAL4>UAS-Luc-CT;*Dm*cry*^{02}/Dm*cry*^{02} flies compared with those in constant darkness (DD) (24.50 h versus 23.75 h, respectively, $P = 0.019$) revealing the BL sensitivity of *Dm*CRY-CT. As with the GFP–CT construct[17], Luc–CT exposure to an MF (300 µT, 3 Hz) was sufficient to shorten the free-running circadian period in the MF-exposed but not the sham group (0 µT MF; Fig. 1c (left); before exposure/after exposure × sham/MF interaction, $F_{1,377} = 7.6$, $P = 0.006$; Extended Data Table 1a,b). Notably, the Luc–CT fusion supports an MF-mediated period shortening compared with the sham group at 50 µT, which is around the strength of the geoMF (Extended Data Table 2). Expression of Luc–CT in the aCC neuron supported a BL-induced increase in action potential firing (1.4-fold), which was increased further in the co-presence of an MF (100 mT; twofold, $P = 0.002$; Fig. 1c (right) and Extended Data Fig. 1b). In summary, these data collectively show that the Luc–CT alone is sufficient to support magnetosensitivity in both circadian and electrophysiological phenotypes. We predict that it does so through its well-described interaction with the redox-sensitive K^+ channel β-subunit HYPERKINETIC (HK)[30]. To improve sensitivity when measuring at the single-neuron level, we chose to use a 100 mT moderate field exposure, using permanent magnets, to saturate the Zeeman effect on the radical pair. According to the RPM[31,32], this is likely to elicit a larger response to MF than exposure to a µT field. We

acknowledge that this field strength is greater than the geoMF; however, the observed effects at 100 mT remain consistent with the RPM and our period-shortening assay showed effects at geoMF strength (Extended Data Fig. 2 and Extended Data Table 2a–c). Control data for electrophysiology are reported in Extended Data Figs. 3 and 4 and Extended Data Table 3.

Fig. 1: Luc–CT is sufficient to support magnetosensitivity.



a
elav-GAL4>UAS-Dmcry;Dmcry02/Dmcry03
BL
APs in 15 s
APs in 15 s
Magnet
Magnet
elav-GAL4>UAS-Dmcry;Dmcry02/Dmcry03
BL + MF
b
elav-GAL4>UAS-Dmcry; Dmcry02/Dmcry03
FFon/FFoff
BL
BL + MF
c
tim-GAL4>UAS-Luc-CT; Dmcry02/Dmcry02
300 μT
Period (h)
DD
BL pre-sham
BL + sham
BL pre-MF
BL + MF
elav-GAL4>UAS-Luc-CT; Dmcry02/Dmcry03
FFon/FFoff
BL
BL + MF
d
tim-GAL4>UAS-Luc-CTW536F; Dmcry02/Dmcry02
300 μT
Period (h)
DD
BL pre-sham
BL + sham
BL pre-MF
BL + MF
elav-GAL4>UAS-Luc-CTW536F; Dmcry02/Dmcry03
FFon/FFoff
BL
BL + MF
Firing in aCC neurons
Circadian period

a, The electrophysiology set-up (permanent magnets are shown in red) (left). Top right, BL-exposure of aCC neurons expressing Dm*cry* increases action potential firing. Bottom right, the co-presence of MF (100 mT) potentiates the effect. The traces are from different preparations. APs, action potentials. **b**, The relative firing frequency (FF) of aCC neurons expressing Dm*cry*. BL increases firing 1.69-fold ($t_9 = 7.72$, $P \leq 0.0001$, $n = 10$, FF_{on}/FF_{off}) compared with in the dark (dashed line). External MF (BL + MF, 100 mT) potentiates the effect to 2.41-fold (BL versus BL + MF, $t_{18} = 3.2$, $P = 0.005$, $n = 10$; Extended Data Fig. 1a). **c**, *tim-GAL4>UAS-Luc-CT;*Dm*cry*^{02}/Dm*cry*^{02} shows period shortening under MF (left) (sham/MF × before/after exposure interaction ($F_{1,377} = 7.6$, $P = 0.006$, three-way ANOVA). $n = 28$ (DD), $n = 108$ (BL pre-sham), $n = 93$ (BL + sham), $n = 104$ (BL pre-MF), $n = 90$ (BL + MF). MF-exposed flies show a significantly shorter period. Four repeats showed the same period shortening under an MF (Extended Data Table 1). Right, Luc–CT supports BL-induced firing (1.4-fold, $t_9 = 4.01$, $P = 0.003$, $n = 10$; Extended Data Fig. 1b) potentiated twofold after BL + MF treatment (BL versus BL + MF, $t_{18} = 3.71$, $P = 0.002$, $n = 10$). **d**, Luc–CT(W536F) revealed significant period shortening after exposure to an MF (left) (significant before/after exposure × MF/sham interaction, $F_{1,198} = 5.1$, $P = 0.025$, two-way ANOVA). $n = 29$ (DD), $n = 47$ (BL pre-sham), $n = 51$ (BL + sham), $n = 52$ (BL pre-MF), $n = 52$ (BL + MF). Post hoc tests are shown in Extended Data Table 1. Right, aCC neurons expressing Luc–CT(W536F) show a twofold change in BL-induced firing ($t_{19} = 6.06$, $P \leq 0.0001$, $n = 20$). The response to BL + MF was variable, but greater than BL alone (2.69-fold, two-way ANOVA, replicates as factor, $F_{1,16} = 5.09$, $P = 0.03$, $n = 20$; Extended Data Fig. 1c). Controls are reported in Extended Data Figs. 3 and 4 and Extended Data Table 3. For FF_{on}/FF_{off} data, the blue asterisks represent significance comparing before versus during BL exposure (same cells, paired two-tailed *t*-tests) and the black asterisks represent comparisons of BL versus BL + MF (different cells, unpaired two-tailed *t*-tests). The reported *n* value for each electrophysiological recording is derived from independent cells from biologically independent animals. The reported *n* values for each circadian period derives from biologically independent animals. Data are mean ± s.e.m. *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$.

Trp536 is probably not an RP partner

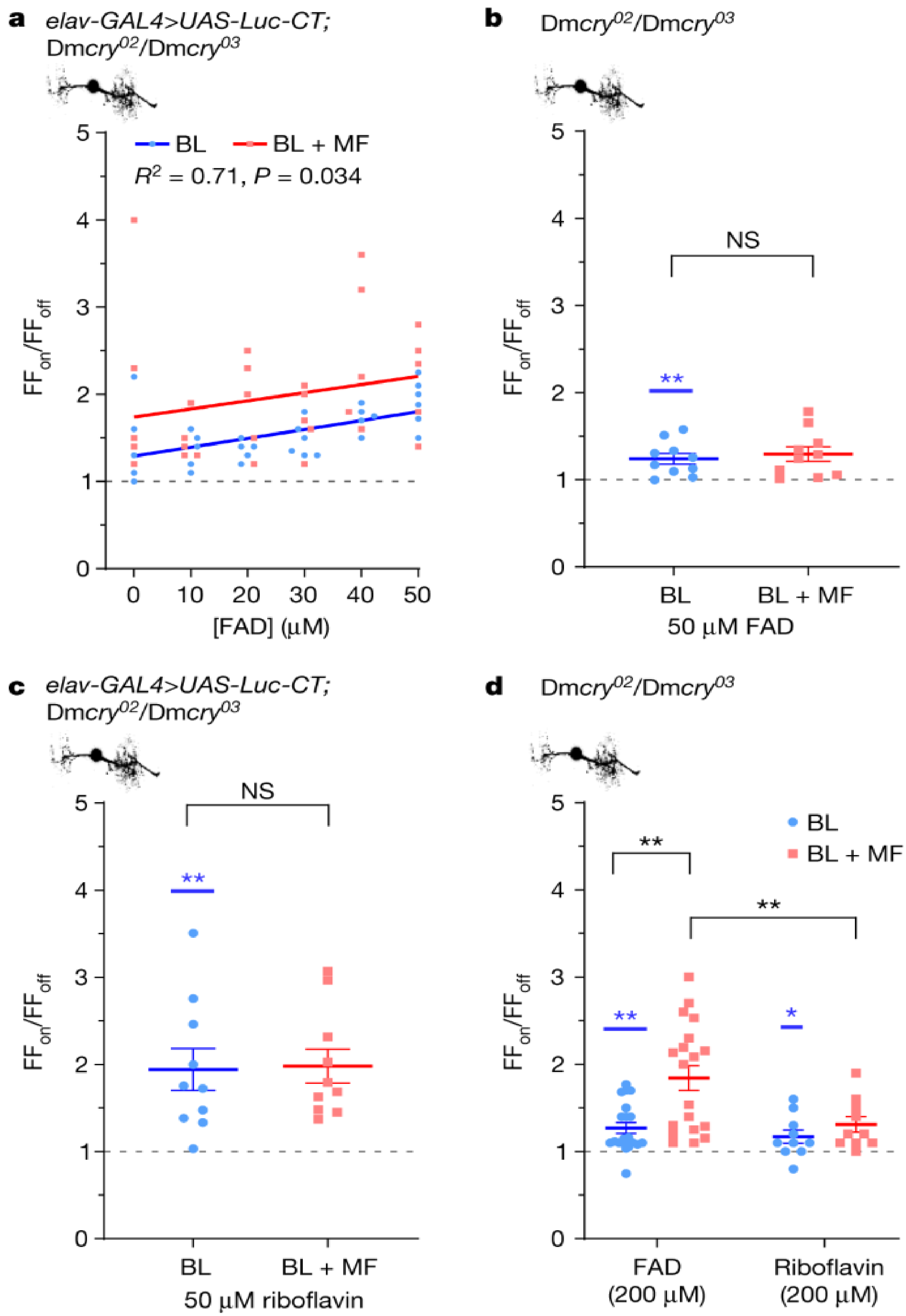
Although CRY-bound FAD may be dispensable, it is possible that FAD in proximity could interact by forming an RP with the sole (non-canonical) Trp in the CT of *Dm*CRY (Fig. 3a). This alternative mechanism may explain why mutations of single Trp residues that constitute the Trp-tetrad are not entirely detrimental to CRY-dependent magnetoreception[3,4,26,33]. The Trp residue in *Dm*CRY-CT has not been implicated in the canonical RPM. However, it is theoretically capable of generating a RP with free FAD reminiscent of the interaction between Flavin mononucleotide and the surface Trp of lysozyme[34]. Thus, we substituted this residue for a redox-inactive phenylalanine (that is, W536F).

Expression of Luc–CT(W536F) was sufficient to lengthen the circadian period in BL versus DD (24.47 h versus 23.95 h, respectively, $P = 0.0017$) indicating that it supports circadian light responsiveness. A two-way analysis of variance (ANOVA) revealed a significant interaction between pre-exposure/exposure and MF/sham treatment ($F_{1,198} = 5.1$, $P = 0.025$; Fig. 1d (left) and Extended Data Table 1c). Expression of this variant also shortened the free-running circadian period when exposed to an MF (300 µT, 3 Hz) compared with pre-exposure and to sham-exposed flies ($P = 0.023$ and $P = 0.015$, respectively, Fisher least significant difference (LSD) test; the stringent Newman–Keuls test narrowly missed the significance threshold for both comparisons ($P = 0.063$ and $P = 0.074$, respectively); Extended Data Table 1d,e). Expression of Luc–CT(W536F) also supported a strong (twofold) BL response on action potential firing in aCC neurons and again a significant, albeit more variable, potentiation under the BL + MF condition (Fig. 1d (right); 2.69-fold, $P = 0.03$, two-way ANOVA replicates as a factor; Extended Data Fig. 1c). The fact that Luc–CT(W536F) does not abolish an MF response argues against an important role for a hypothetical RP between Trp536 and FAD. Indeed, the weaker MF response may be structural in origin[35]. An arginine (Arg532) in close proximity may form a cation-π interaction with Trp536 to stabilize an alpha-helical conformation[36] that would be disrupted by the W536F substitution, yet the MF effect is still detectable.

Free FAD supports magnetosensitivity

The fact that Luc–CT(W536F) is sufficient to support magnetosensitivity implies that a different, non-CRY, RP is involved. In this regard, it is notable that free FAD has the ability to generate a magnetically sensitive RP through intramolecular electron transfer[7,37]. To investigate this, we supplemented additional FAD to aCC neurons through the internal patch saline. Increasing FAD (range, 10 to 50 μM in the patch pipette) potentiates the efficacy of Luc–CT to mediate BL-dependent increases in action potential firing (Fig. 2a; $R^2 = 0.71$, $P = 0.034$), an effect that is enhanced in the presence of BL + MF (100 mT, $P = 0.015$). Notably, MF potentiation is by a fixed proportion relative to BL at each FAD concentration tested (evidenced by equal gradients of lines of best fit). This is a prediction of the RPM; provided that biological saturation is not limiting, the proportional magnetically induced change should remain constant[31]. In the absence of Luc–CT, FAD (up to 50 μM) induced a weak, but significant, response to BL ($P = 0.03$; Fig. 2b); however, no potentiating effect was observed in the BL + MF condition (Extended Data Fig. 5a).

Fig. 2: Free FAD potentiates the effect of Luc–CT and, at high concentration, supports magnetosensitivity alone.



a
elav-GAL4>UAS-Luc-CT;
Dmcry02/Dmcry03
BL
BL + MF
R2 = 0.71, P = 0.034
FFon/FFoff
[FAD] (μM)
b
Dmcry02/Dmcry03
NS
**
BL
BL + MF
50 μM FAD
c
elav-GAL4>UAS-Luc-CT;
Dmcry02/Dmcry03
NS
**
BL
BL + MF
50 μM riboflavin
d
Dmcry02/Dmcry03
BL
BL + MF
**
**
*
FAD
(200 μM)
Riboflavin
(200 μM)

a, Exposing aCC neurons expressing Luc–CT to FAD (through the recording pipette) increases the response to BL ($R^2 = 0.71$, $F_{1,4} = 10.1$, $P = 0.03$, linear regression). The co-presence of an MF (100 mT) potentiates the response ($F_{1,9} = 9.06$, $P = 0.015$, analysis of covariance (ANCOVA) model). $n = 5$ except for BL 30 μM and 50 μM FAD, for which $n = 6$. **b**, The addition of FAD (50 μM) supports BL sensitivity (Extended Data Fig. 5a), but not magnetosensitivity in the absence of Luc–CT ($t_{18} = 0.521$, $P = 0.609$, $n = 10$). **c**, The addition of riboflavin (50 μM) to aCC neurons expressing Luc–CT, supports the response to BL, but not MF potentiation ($t_{18} = 0.12$, $P = 0.91$). $n = 10$. **d**, Increased FAD (200 μM) in the Dm*cry*^{02}-null background supports a BL-induced change in firing (1.27-fold, $t_{19} = 4.29$, $P = 0.0004$, $n = 20$; Extended Data Fig. 5c). The co-presence of an MF ($n = 19$) significantly potentiates this effect (1.84-fold, two-way ANOVA, $F_{1,55} = 3.51$, $P = 0.066$, Newman–Keuls post hoc $P = 0.003$). Riboflavin (200 μM) shows a similar BL effect (1.17-fold, $t_9 = 2.33$, $P = 0.045$, $n = 10$; Extended Data Fig. 5c) but no MF potentiation (1.31-fold, Newman–Keuls post hoc, $P = 0.67$, $n = 10$). Raw data are reported in Extended Data Fig. 5. The blue asterisks represent significance values before versus during BL exposure (same cells, paired two-tailed *t*-tests) and the black asterisks represent comparisons of the BL versus BL + MF condition (different cells, unpaired two-tailed *t*-tests). NS, not significant. The reported *n* value for each electrophysiological recording is derived from independent cells from biologically independent animals. Data are mean ± s.e.m. NS, $P \geq 0.06$.

FAD autoreduction occurs after electron transfer from the adenine side chain to the photoexcited isoalloxazine, generating an intramolecular RP[7,38]. To test whether this autoreduction is likely to support magnetosensitivity, we introduced riboflavin to the cell through the internal patch saline. Although riboflavin contains the same isoalloxazine chromophore and can populate photoexcited triplet states[39], it lacks an adenine diphosphate side chain (Extended Data Fig. 6) and is therefore unable to generate the same intramolecular RP[40]. Riboflavin (50 μM) in the presence of Luc–CT supported a BL effect (~1.94-fold; Fig. 2c), but there was no additional increase under the BL + MF condition (100 mT; $P = 0.9$; Fig. 2c and Extended Data Fig. 5b).

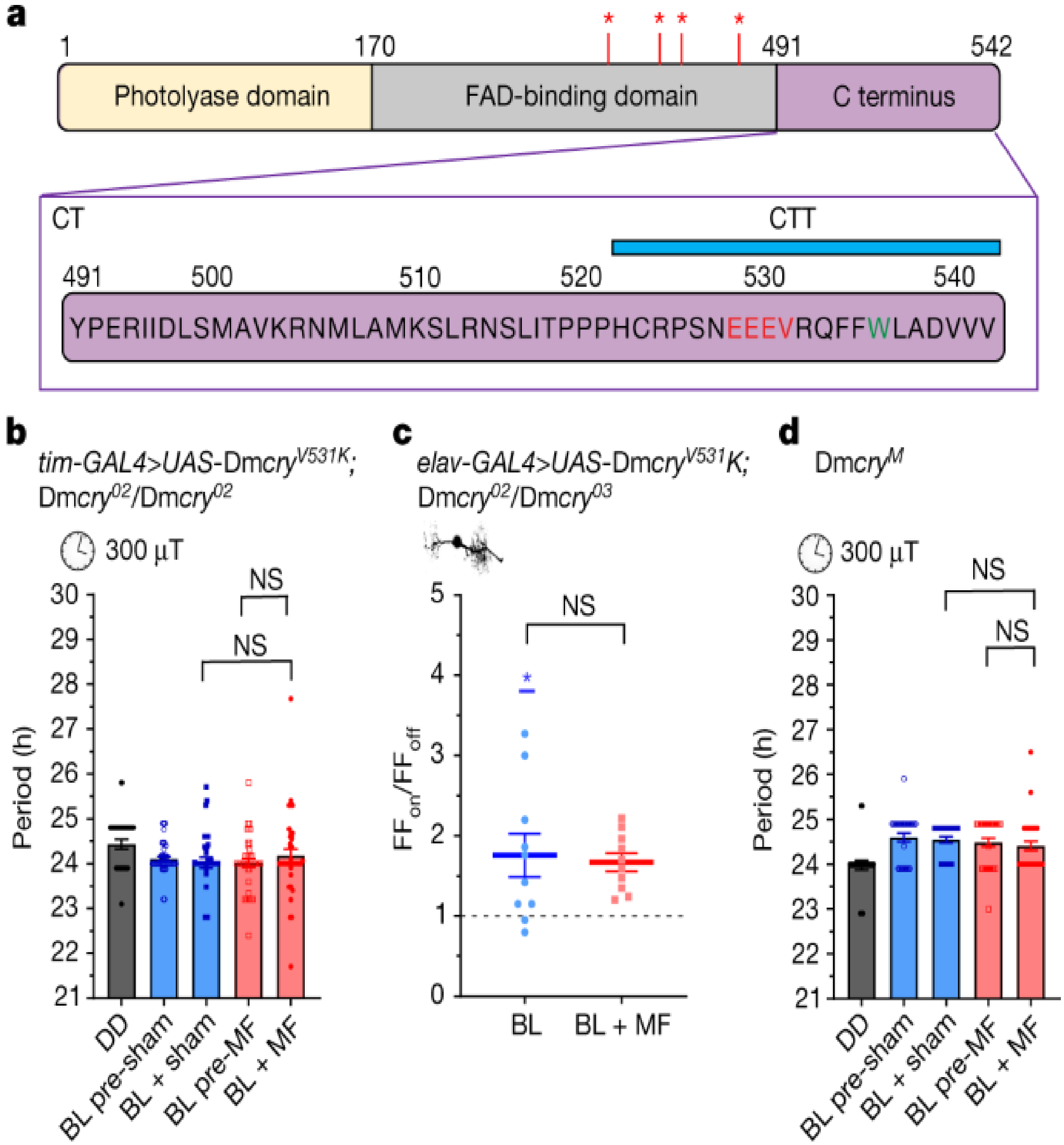
Our results are consistent with an interaction between FAD and Luc–CT, possibly in a complex with other, unknown, molecules, which together may facilitate the transduction of a magnetic field. Furthermore, our data suggest that molecules other than CRY are able to generate magnetically sensitive RPs and produce a biological effect under appropriate conditions. In vitro spectroscopy has shown that BL-photoexcited FAD generates RPs that are responsive to MFs[41], and it appears probable that FAD is responsible for MF effects that were recently observed on cellular autofluorescence[37]. Thus, FAD (but not riboflavin) at higher concentrations may act as a magnetoreceptor. To test this, we recorded from aCC neurons in the Dm*cry*-null background, which shows no overall BL or MF response (Extended Data Fig. 4d). We observed that high levels of FAD in the internal patch saline (200 µM) were sufficient to support a BL-dependent increase in action potential firing without the need for Luc–CT (1.27-fold; Fig. 2d and Extended Data Fig. 5c). Notably, this effect was potentiated in the presence of an MF (100 mT, 1.84-fold, $P = 0.003$; Fig. 2d). Cells supplemented with riboflavin (200 µM) showed an increase in action potential firing in response to BL (Fig. 2d and Extended Data Fig. 5c) but did not show potentiation of the response in an MF (100 mT, $P = 0.67$). The fact that high levels of FAD alone are sufficient to support magnetosensitivity suggests that *Dm*CRY-CT acts as an adaptor protein, bringing photoactivated FAD close to effectors, possibly HK. Proximity may enable HK to be activated directly by the resultant change in oxidative state that results from the photoactivation of FAD. Very high levels of FAD negate this requirement. In the presence of *Dm*CRY-CT, the amount of photoactivated FAD required is presumably lower and, therefore, more reflective of normal physiological amounts of this flavin.

The integrity of the CTT is important

The less-robust response from Luc–CT(W536F) (Fig. 1d) suggests that the integrity of the CTT might be important to its role in facilitating magnetosensitivity. The CTT of *Dm*CRY also contains several linear motifs, including putative PDZ-binding sequences[42] (for example, EEEV 528–531; Fig. 3a). PDZ proteins function as modular scaffolds that direct the cellular localization of signalling molecules, such as ion channels (for example, Shaker K^+)[43,44], and the assembly of signalling partners (including *Dm*CRY)

into a ‘signalplex’ of the phototransduction cascade in the *Drosophila* eye[21,22]. To examine the importance of the CTT structure and, specifically, to determine whether the putative PDZ-binding motif at residues 528–531 regulates magnetosensitivity, we mutated valine to lysine at position 531 (V531K)[42] in full-length *Dm*CRY. Pan-circadian expression (that is, using the *tim-GAL4* driver) of Dm*cry*^{V531K} in a Dm*cry*^{02}-null background retained circadian light sensitivity with slight period shortening (24.51 h (DD) versus 24.20 h (BL), $P = 0.005$, Grubbs outlier test excluded a single very weakly rhythmic short period (20.3 h) fly in DD; Fig. 3b and Methods), but did not support a whole-organism behavioural response to MF (interaction, $F_{1,158} = 0.55$, $P = 0.33$; Fig. 3b and Extended Data Table 4a). Expression of Dm*cry*^{V531K} in aCC neurons showed the expected effect of BL on action potential firing (1.76-fold; Fig. 3c and Extended Data Fig. 7a). As in the circadian assay, this variant was unable to support magnetosensitivity in the aCC neuron (100 mT, $P = 0.77$; Fig. 3c). The loss of an MF effect, but the retention of a BL response for *Dm*CRY(V531K), is reminiscent of the MF phenotype of the UAS-*Dm*CRYΔ variant[3,4], which carries a deletion of the CTT[16]. To further validate this result, we used Dm*cry*^{M}—a variant lacking the final 19 amino acids of the CTT[45], including the putative PDZ domain centred around residue 531. Expression of Dm*cry*^{M} retained circadian light sensitivity but did not show MF-induced period shortening as tested at both 300 µT (3 Hz, $F_{1,122} = 0.021$, $P = 0.89$; Fig. 3d) and 50 µT (3 Hz, $F_{1,180} = 0.3$, $P = 0.6$; Extended Data Fig. 7b and Extended Data Table 4b–d). These results confirm the CTT as a probable mediator of the MF response, where it may serve to facilitate the formation of protein complexes that transduce the magnetic signal.

Fig. 3: Integrity of the CTT is required for it to facilitate magnetosensitivity.



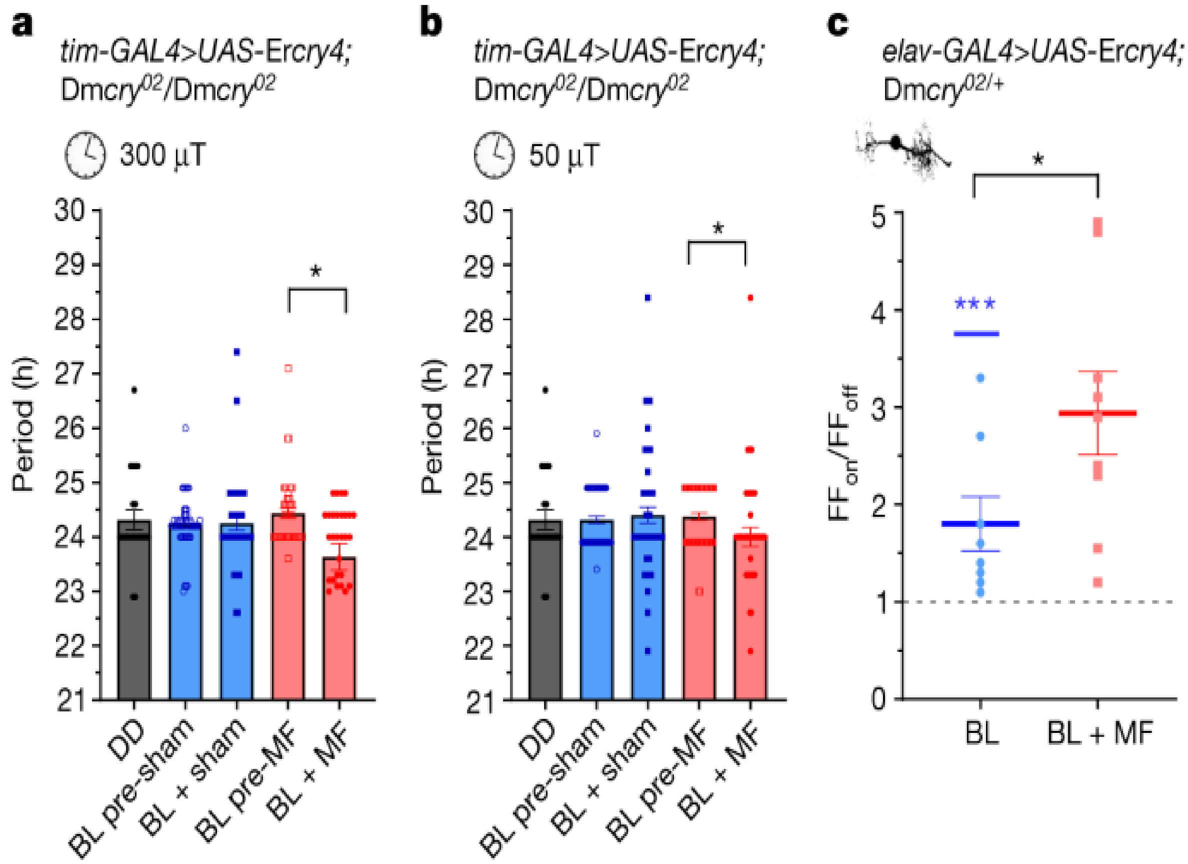
a, Schematic of the domain structure of full-length *Dm*CRY, including the CT (amino acids 491–542) and CTT (amino acids 521–542). The four Trp residues, presumed to be essential for the canonical RPM, are indicated by red asterisks. A putative PDZ-binding site (EEEV 528–531, shown in red) was mutated (Val531) *Dm*CRY(V531K). The Trp residue (Trp536) mutated in Luc–CT(W536F) is shown in green. **b**, Dm*cry*^{V531K} expressed in clock neurons (*tim-GAL4*) does not support magnetosensitivity in the circadian period-shortening assay. A two-way ANOVA revealed no significant main effects or interaction effects (interaction, $F_{1,52} = 0.09$, $P = 0.77$). $n = 53$

(DD), $n = 43$ (BL pre-sham), $n = 41$ (BL + sham), $n = 41$ (BL pre-MF), $n = 38$ (BL + MF) (Extended Data Table 4a). **c**, Expression of Dm*cry*^{V531K} in aCC neurons is sufficient to support BL sensitivity ($t_9 = 2.934$, $P = 0.017$, $n = 10$; Extended Data Fig. 7a) but not potentiation in the BL + MF condition (100 mT, $t_{18} = 0.299$, $P = 0.768$, $n = 10$). **d**, Expression of Dm*cry*^{M}, a truncated CRY variant lacking the terminal 19 amino acids, including the PDZ-binding motif (528–531), does not support sensitivity to a 300 µT MF (3 Hz, two-way ANOVA, $F_{1,122} = 0.021$, $P = 0.89$). $n = 26$ (DD), $n = 31$ (BL pre-sham), $n = 31$ (BL + sham), $n = 32$ (BL pre-MF), $n = 32$ (BL + MF) (Extended Data Table 4b). The blue asterisks represent significance values before versus during BL exposure (same cells, two-tailed paired *t*-tests) and the black asterisks represent comparisons of the BL versus BL + MF condition (different cells, unpaired two-tailed *t*-tests). The reported *n* value for each electrophysiological recording is derived from independent cells from biologically independent animals. The reported *n* values for each circadian period derives from biologically independent animals. Data are mean ± s.e.m.

*Er*CRY4 supports magnetosensitivity

Recent in vitro spectroscopy studies have suggested that CRY4, encoded in the genome of the European robin, *Erithacus rubecula*, a migratory songbird[46], may represent the magnetoreceptor that is responsible for long-distance navigation in this species. We generated a *UAS*-Er*cry4* transgene and expressed it in clock neurons of the fly. We observed significant period-shortening on exposure to an MF compared with sham treatment at 300 µT MF (3 Hz, two-way ANOVA $F_{1,238} = 4.4$, $P = 0.036$; Fig. 4a) or 50 µT (3 Hz, two-way ANOVA, interaction, $F_{1,237} = 3.97$, $P = 0.047$; Fig. 4b and Extended Data Table 5). Expression of *E. rubecula cry4* (Er*cry4*) in aCC neurons was also sufficient to render the cell sensitive to BL (1.8-fold) and substantially sensitive to an external MF (100 mT, 2.94-fold, $P = 0.046$; Fig. 4c and Extended Data Fig. 7c).

Fig. 4: *Er*CRY4 is sufficient to support MF sensitivity in *Drosophila*.



a, Expression of Er*cry4* in *Drosophila* clock neurons (through *tim-GAL4*) results in significant period shortening in the presence of a 300 µT MF (3 Hz, two-way ANOVA interaction, $F_{1,238} = 4.4$, $P = 0.036$). $n = 22$ (DD), $n = 49$ (BL pre-sham), $n = 41$ (BL + sham), $n = 49$ (BL pre-MF), $n = 30$ (BL + MF) (Extended Data Table 5a–c). **b**, Period shortening is also present in a 50 µT MF (3 Hz, two-way ANOVA interaction, $F_{1,237} = 3.97$, $P = 0.047$). $n = 22$ (DD), $n = 64$ (BL pre-sham), $n = 59$ (BL + sham), $n = 64$ (BL pre-MF), $n = 54$ (BL + MF) (Extended Data Table 5d,e). **c**, Relative firing-frequency recordings of aCC motorneurons expressing Er*cry4* for the BL versus BL + MF condition. BL exposure increases action potential firing (1.8-fold, $t_7 = 3.6$, $P = 0.0088$, $n = 8$; Extended Data Fig. 7c), an effect that is potentiated by the co-presence of an MF (100 mT, 2.94-fold, BL versus BL + MF, $t_{15} = 2.17$, $P = 0.046$, $n = 9$). The blue asterisks represent significance values for before versus during BL exposure (same cells, paired two-tailed *t*-tests) and the black asterisks represent comparisons of BL versus BL + MF (different cells, unpaired two-tailed *t*-tests). The reported *n* value for each electrophysiological recording is derived from independent cells from

biologically independent animals. The reported *n* value for each circadian period derives from biologically independent animals. Data are mean ± s.e.m.

Conclusions

We have observed that, contrary to several reports[2,14], but not others[3], full-length *Dm*CRY is sufficient, but not strictly necessary, to mediate magnetosensitivity. The expression of the C-terminal 52 residues of *Dm*CRY is sufficient to support magnetosensitivity in both single-neuron and whole-animal assays. Our results challenge the canonical CRY-dependent RPM model of animal magnetoreception (based on the requirement for full-length CRY, including FAD binding and the Trp chain), yet are consistent with an RPM. Although it remains unclear whether Luc–CT binds to FAD directly, the Luc–CT response is potentiated by increasing the cytosolic availability of FAD—a common biological redox cofactor—implying that redox reactions are at the core of magnetosensitivity[47]. We cannot exclude that alternative RPs that are not directly photochemically generated may also contribute to magnetoreception, which would be consistent with a growing list of examples reporting RP-mediated magnetoreception in the dark[30,48,49,50]. The synergistic interaction between Luc–CT and free FAD argues that the former facilitates the formation of a complex that enables the transduction of a magnetically derived signal by the latter. Moreover, free FAD itself can mediate a magnetic response in vivo but at high, non-physiological, levels. We interpret these results to suggest that evolution has shaped the defining element of CRY, its CT, to bring the RP to the proximity of cellular effectors such as HK. Thus, through protein–protein interactions, CRY may potentiate the weak activity of the geoMF on any associated RP. In this regard, the primary role of CRY would be that of a magnetotransducer rather than a magnetoreceptor.

The unexpected observation that robin *Er*CRY4 can also mediate MF effects in *Drosophila*, in both circadian and electrophysiological assays, argues that the fly is an excellent tractable model system to dissect the molecular component of magnetoreception. This has raised the question of why flies have a magnetic sense, given that they do not navigate and migrate in the same way as birds do. Although our circadian phenotype is somewhat

contrived and seeks to use a sensitized CRY background (dim constant BL) to provide the best opportunity for observing any MF effects, *Dm*CRY is not only a circadian photoreceptor—it also mediates geotaxis[51]. Independent studies have revealed that geotaxis shows a *Dm*CRY-dependent magnetosensitivity[52,53]. Notable results have suggested that flies exposed to an MF as embryos are imprinted on the MF in which they develop and, as adults, they prefer to forage with downward movement in their home MF[54]. As *D. melanogaster* is well known to forage/mate/oviposit on rotten fruits that are usually found at the ground level, this geotactic magnetic sense would appear to have fitness value.

In conclusion, our observations suggest an ancient and ubiquitous effect of MFs on biological RPs. Through CRY, evolution has optimized such an effect by bringing together two functions, receptor and transductor, that are required for magnetosensing but not necessarily as parts of the same molecule. The fact that *Drosophila* (and other non-migrating animals) can sense external magnetic fields has been reported by many independent groups[55]. This appears reflective of the physiochemical properties of flavins such as FAD to form RPs. In animals that navigate, this mechanism has presumably been adapted to underpin this behaviour, but the underlying physiochemical properties of CRY-dependent magnetosensitivity appear to be shared across navigating and non-navigating animals.

Methods

Fly stocks

For larval aCC neuron recordings, embryos were raised at 25 °C under a 12 h–12 h light–dark cycle until third instar wall-climbing larvae (L3) emerged, these were then kept in the dark through the day of recording to minimize light-dependent *Dm*CRY degradation. Recordings were conducted between circadian time hours 2–10. Flies were maintained on standard corn meal medium at 25 °C. The driver line $elav^{C155}$*-GAL4;;* Dmcry^{03} was obtained by crossing the driver line from Bloomington Stock centre (BL458) into a Dmcry^{03} background as described previously[56]. *Dmcry*^{02} flies were obtained from Bloomington Stock centre (BL86267). *tim-GAL4;cry*^{02} and *UAS*-

Dm*cry;*Dm*cry*^{02} were described previously[3,57]. Dm*cry*^{M} (supplied by D. Dolezel) has a stop codon inserted at amino acid 523 and lacks the final 19 amino acids of the CTT, which includes the putative PDZ-binding motif at 531[45] (Fig. 3a). The generation of new transgenic flies for this study is described below.

Molecular cloning of Luc–CT

The luciferase coding sequence was cloned from the *UAS*-*Luc*-Dm*cry* fly line[18] and subsequently amplified using the following primers to include overhangs compatible for the NEB Gibson Assembly assay: F, TATCCTTTACTTCAGGCGGCCGCATGGAAGACGCCAAAAACATAAAGAAAGG; R, TCCGGATACTCGAGCACGGCGATCTTTCCGCCC.

The CT portion of Dm*cry* was produced by gene synthesis (GeneArt, Thermo Fisher Scientific) based on the original *GFP*-*CT* construct[3]. *CT* was designed to include 5′ and 3′ overhangs compatible with subsequent NEB Gibson Assembly assay (sequence, AAAGATCGCCGTGCTCGAGTATCCGGAGCGGATCATTGATTTGTCCATGGCCGTGAAGCGCAACATGCTGGCCATGAAGTCCCTGCGCAACAGCCTGATCACCCCCCCACCACATTGCCGCCCCAGCAATGAGGAGGAAGTGCGCCAGTTCTTCTGGCTGGCCGATGTGGTGGTGTAATCTAGAGGATCTTTGTGAAGGA).

The pJFRC2-10XUAS-IVS-mCD8::GFP (Addgene, 26214) plasmid was digested with NotI and XbaI. A Gibson Assembly assay was performed to ligate pJFRC, *Luc* and *CT* in a single-step reaction. A Myc tag was produced by annealing single oligos designed using the Oligator software (https://gcat.davidson.edu/iGem10/) and ligated 5′ of Luc-CT. In brief, 5 µM of the oligos 47-mer Top1, 5′-GATCTCACAATGGAACAGAAGCTGATCTCCGAGGAGGACCTGGGCGC; 47-mer Bottom1, 5′ GGCCGCGCCCAGGTCCTCCTCGGAGATCAGCTTCTGTTCCATTGTGA, resulting in 5′ BglII and 3′ NotI overhangs once annealed, were diluted in 1× annealing buffer (0.1 M NaCl; 10 mM Tris-HCl, pH 7.4), boiled in 500 µl of H_2O for 10 min and left overnight to cool down to room temperature to hybridize. The pJFRC2-10XUAS-Luc-CT plasmid was then digested with

BglII and NotI. The fragment encoding the Myc tag was then ligated using standard methods. After sequence validation, the plasmid was injected into the *y w, M(eGFP,vas-int,dmRFP)ZH-2A;P{CaryP}attP40* (stock 13–40, Cambridge University Fly facility) using the Phi31 integrase system for insertion. The resulting transformants were subsequently backcrossed into the w^{1118} background for seven generations. The Luc–CT(W536F) transgene was generated by gene synthesis (Eurofins) and subcloned into pJFRC-MUH through the 5′ NotI and 3′ XbaI restriction sites. Transgenic injections for Luc–CT(W536F) were carried out by Manchester University Fly Facility using the same *y w M(eGFP,vas-int,dmRFP)ZH-2A;P{CaryP}attP40* line (stock 13–40, University of Cambridge Fly Facility). The Er*cry4* transgenic was generated by gene synthesis (NBS Biological) and subcloned into pJFRC-MUH through the NotI and KpnI restriction sites. Injection was carried out by Bestgene into *y w;PBac{y[+]-attP-3B}VK00002* (Bloomington, 9723) using the Phi31 integrase system.

HA-Dmcry^{V531K} (containing a HA tag at the N terminus of the encoded protein) was already available as a clone in the yeast plasmid pEG202 (ref. [42]). It was released by EcoRI-XhoI digestion and subcloned into pUAST[58]. Transgenics were produced by P-element transformation by the University of Cambridge Fly Facility using the line S(6)1 inserted on chromosome 2.

Electrophysiology

The experimenter was blinded to genotype during both recordings and subsequent data analysis. L3 larvae were dissected under extracellular saline as described[59] with the only modification being a red filter applied to both the dissecting light and compound microscope to minimize *Dm*CRY degradation before and during recordings. Thick-walled borosilicate glass electrodes (GC100F-10; Harvard Apparatus) were fire-polished to resistances of 10–15 MΩ. Recordings were made using the Multiclamp 700B amplifier controlled by pCLAMP (v.10.4) and the Digidata 1440A analogue-to-digital converter (Molecular Devices). Only cells with an input resistance of ≥500 MΩ were used. Traces were filtered at 10 kHz and sampled at 20 kHz. The extracellular saline solution contained the following: 135 mM NaCl, 5 mM KCl, 4 mM $MgCl_2.6H_2O$, 2 mM $CaCl_2.2H_2O$, 5 mM TES and 36 mM sucrose, pH 7.15. The intracellular patch solution contained

the following (in mM): 140 mM potassium-d-gluconate, 2 mM $MgCl_2.6H_2O$, 2 mM EGTA, 5 mM KCl and 20 mM HEPES, pH 7.4. KCl and $CaCl_2$ were from Thermo Fisher Scientific; sucrose was from BDH; all of the remaining chemicals were from Sigma-Aldrich. Mecamylamine (1 mM) was applied to all preparations to isolate the aCC motoneurons from excitatory cholinergic synaptic input. For recordings supplemented with additional FAD or riboflavin (Sigma-Aldrich), dilutions were made up in intracellular saline and kept in the dark.

Photoactivation and magnetic field application

Light stimulation was supplied by a blue LED (470 nm, Cairn Research) at a power of around 2.2 mW cm^{-2}, a value used previously to stimulate *Dm*CRY[60]. Each cell was injected with a variable amount of constant current until the threshold potential was reached and the neuron was allowed to settle, for some minutes, until action potential firing was stable at ~5–7 Hz. Once a stable firing rate was achieved, each neuron was recorded for at least 20 s before exposure to BL illumination for 30 s. No significant change to the action potential firing rate was observed without BL illumination. Magnetic exposure was provided by two NdFeB static magnets mounted around the preparation at a distance that provided an MF of 100 mT (±5 mT). Field strength was measured using the 5180 Gauss/Tesla Meter (F.W. Bell). This method is essentially identical to that used previously[4].

Statistical analysis of electrophysiological recordings

The D'Agostino–Pearson analysis showed our data to be normally distributed and parametric tests were therefore applied in all cases. Data are shown as mean ± s.e.m. To determine BL sensitivity, paired two-tailed *t*-tests were used to compare the number of action potentials that a neuron fired in the 15 s after light stimulation versus the number of action potentials in the 15 s before light exposure. For a comparison between BL and BL + MF, the number of action potentials in the 15 s before and after BL or BL + MF exposure was used to determine the firing fold change (FF_{on}/FF_{off}) for each cell. Statistical significance for MF potentiation against the BL effect alone was determined using unpaired two-tailed *t*-tests to compare the firing fold

change for the BL dataset versus the BL + MF dataset. In cases in which multiple genotypes/conditions were tested simultaneously, two-way ANOVA with Newman–Keuls post hoc testing was used. For the FAD dose–response curve, a BL effect of [FAD] was determined by linear regression fitting and significance was determined using an ANCOVA model. Average MF potentiation of the Luc–CT FAD dose–response (Fig. 2a) was determined on the basis of the intercept of the *y* axis. Unpaired two-tailed *t*-tests were also applied in the Extended Data (Figs. 1a–c, 3b,c, 4a–d, 5a–c and 7a,c) to compare the number of action potentials in the before BL ± MF conditions, as well as to BL and BL + MF exposures. Control lines were also compared to their respective experimental genotype by both one-way ANOVA (with BL and BL + MF recordings separated) and by two-way ANOVA. Raw data are reported in the Extended Data (Figs. 1a–c, 3b,c, 4a–d and 7a,c).

Behavioural analyses and statistics

Circadian locomotor activity was recorded using a *Drosophila* Trikinetics Monitor 2 (ref. [3]). To test the effects of MF on the free-running circadian period of locomotor activity, we used a modified version of the Schuderer apparatus, which consists of two independent double-wrapped coils placed inside two mu-metal boxes within a commercial incubator. The shielded four quadratic Helmholtz coil systems produce a homogenous, linearly polarized, *B* field with perpendicular orientation to the horizontal plane of the Trikinetics monitors. Each coil is formed with a pair of wires, with the current passing in the same direction through both wires for MF exposure but in opposite directions to provide a sham exposure condition (0 T). A computer randomly assigns the MF- and sham-exposed chambers and the experiment is performed in a blinded manner[17].

Flies (aged 1–3 days) were first entrained at 20 °C in the apparatus under a dim BL–DD 12 h–12 h cycle for 3 days, before being pre-exposed to continuous BL for 7 days, followed by exposure to BL + MF or BL + sham for a further 7 days[3]. Thus, there were four measurements, the pre-exposure (BL) period of flies that were subjected to an MF or sham, plus the exposure period for both (BL + MF and BL + sham). A fifth control condition examined the period of *Luc-CT;*Dm*cry*^{02} flies in DD without exposure. All of the experiments were performed using a low-frequency 3 Hz field at 300

μT or 50 μT and dim BL at 0.15–0.25 μW cm^{-2}, wavelength 450 nm, 40 nm broad range (RS Components). The driver *tim-GAL4* was used to express UAS-Dm*cry* transgenes as previously described[3]. Locomotor data were collected in 1 min time bins but analysed in 10 and 30 min bins.

Rhythmicity and period were determined using spectral analysis using a MATLAB-based version of the BeFly program[60]. Statistical analysis of period was performed using ANOVA with either Statistica (Statsoft) for factorial analyses or Prism (GraphPad) for one-way ANOVA. Although there was a clear prediction that Luc–CT flies would have a shorter period under an MF[3], we nevertheless used the stringent Newman–Keuls post hoc test to compare groups after factorial ANOVA with the more liberal Fisher LSD test. To compare the DD periods with those from the BL pre-exposure conditions, we used unpaired *t*-tests. Circadian data were first tested using a Grubb's outlier test (GraphPad Prism, alpha = 0.01, two-sided, $Z = 5.3$). One datum from the DD data of *Dm*CRY(V531K) that represented the least-robust single period in the dataset with an anomalous period of 20.3 h (8 s.d. away from the mean) was identified and removed.

Determination of sample size

Sample size for electrophysiological recordings was based on preliminary experiments as well as published work[4], which showed reproducible effects. For weaker effects, sample size was increased. For circadian work, sample size was based on previous experience and extensive published work[3] as well as a power calculation using the s.d. of preliminary experiments. This indicated, with a power of 80% at a 95% confidence, that we could reliably detect a 0.4 h difference in period with an $n = 16$, and a 0.3 h difference with an $n = 28$. All comparisons of sham and electromagnetic field exposure before and after had an $n > 28$.

Replication

Electrophysiological experiments were repeated twice with the following exceptions: (1) for Fig. 2a, the FAD dose–response curve due the already high number of recordings for the overall experiment ($n = 62$); (2) for Fig. 2c, riboflavin (50 μM) manipulation in the presence of Luc–CT; and (3) for

Fig. 2b,d, 50 µM FAD and 200 µM exposure in the Dm*cry*-null background (however, the effect size for 200 µM FAD (Fig. 2d) was increased to $n = 20$) and (4) for *Er*CRY4, in which the effect was clear and consistent as to not require further laborious recordings. Circadian MF experiments were repeated at least twice.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The analysed data are included in Extended Data Figs. 1–7 and Extended Data Tables 1–5. All raw data for circadian locomotor experiments have been deposited in the Open Science Forum repository (https://osf.io/6fnra/?view_only=1b825d853813402f9aa41927e5c3cc0f). Raw data for the electrophysiology experiments are available on request.

References

1. Wiltschko, W. & Wiltschko, R. Magnetic orientation and magnetoreception in birds and other animals. *J. Comp. Physiol. A* **191**, 675–693 (2005).

2. Hore, P. J. & Mouritsen, H. The radical-pair mechanism of magnetoreception. *Annu. Rev. Biophys.* **45**, 299–344 (2016).

3. Fedele, G. et al. Genetic analysis of circadian responses to low frequency electromagnetic fields in *Drosophila melanogaster*. *PLoS Genet.* **10**, e1004804 (2014).

4. Giachello, C. N. G., Scrutton, N. S., Jones, A. R. & Baines, R. A. Magnetic fields modulate blue-light-dependent regulation of neuronal firing by cryptochrome. *J. Neurosci.* **36**, 10742–10749 (2016).

5. Schwarze, S. et al. Weak broadband electromagnetic fields are more disruptive to magnetic compass orientation in a night-migratory songbird (*Erithacus rubecula*) than strong narrow-band fields. *Front. Behav. Neurosci.* **10**, 55 (2016).

6. Günther, A. et al. Double-cone localization and seasonal expression pattern suggest a role in magnetoreception for European robin cryptochrome 4. *Curr. Biol.* **28**, 211–223 (2018).

7. Antill, L. M. & Woodward, J. R. Flavin adenine dinucleotide photochemistry is magnetic field sensitive at physiological pH. *J. Phys. Chem. Lett.* **9**, 2691–2696 (2018).

8. Wan, G., Hayden, A. N., Iiams, S. E. & Merlin, C. Cryptochrome 1 mediates light-dependent inclination magnetosensing in monarch butterflies. *Nat. Commun.* **12**, 771 (2021).

9. Kyriacou, C. P. & Rosato, E. Genetic analysis of cryptochrome in insect magnetosensitivity. *Front. Physiol.* **13**, 1522 (2022).

10. Ritz, T., Adem, S. & Schulten, K. A model for photoreceptor-based magnetoreception in birds. *Biophys. J.* **78**, 707–718 (2000).

11. Ritz, T. et al. Magnetic compass of birds is based on a molecule with optimal directional sensitivity. *Biophys. J.* **96**, 3451–3457 (2009).

12. Kao, Y. T. et al. Ultrafast dynamics and anionic active states of the flavin cofactor in cryptochrome and photolyase. *J. Am. Chem. Soc.* **130**, 7695–7701 (2008).

13. Immeln, D., Weigel, A., Kottke, T. & Pérez Lustres, J. L. Primary events in the blue light sensor plant cryptochrome: intraprotein electron and proton transfer revealed by femtosecond spectroscopy. *J. Am. Chem. Soc.* **134**, 12536–12546 (2012).

14. Rodgers, C. T. & Hore, P. J. Chemical magnetoreception in birds: the radical pair mechanism. *Proc. Natl Acad. Sci. USA* **106**, 353–360 (2009).

15. Giovani, B., Byrdin, M., Ahmad, M. & Brettel, K. Light-induced electron transfer in a cryptochrome blue-light photoreceptor. *Nat. Struct. Biol.* **10**, 489–490 (2003).

16. Rosato, E. et al. Light-dependent interaction between *Drosophila* CRY and the clock protein PER mediated by the carboxy terminus of CRY. *Curr. Biol.* **11**, 909–917 (2001).

17. Dissel, S. et al. A constitutively active cryptochrome in *Drosophila melanogaster*. *Nat. Neurosci.* **7**, 834–840 (2004).

18. Peschel, N., Chen, K. F., Szabo, G. & Stanewsky, R. Light-dependent interactions between the *Drosophila* circadian clock factors cryptochrome, jetlag, and timeless. *Curr. Biol.* **19**, 241–247 (2009).

19. Zoltowski, B. D. et al. Structure of full-length *Drosophila* cryptochrome. *Nature* **480**, 396–399 (2011).

20. Levy, C. et al. Updated structure of *Drosophila* cryptochrome. *Nature* **495**, 396–399 (2013).

21. Ye, F. et al. An unexpected INAD PDZ tandem-mediated plcβ binding in *Drosophila* photo receptors. *eLife* **7**, e41848 (2018).

22. Mazzotta, G. M. et al. Calmodulin enhances cryptochrome binding to INAD in *Drosophila* photoreceptors. *Front. Mol. Neurosci.* **11**, 280 (2018).

23. Müller, P., Brettel, K., Grama, L., Nyitrai, M. & Lukacs, A. Photochemistry of wild-type and N378D mutant *E. coli* DNA photolyase with oxidized FAD cofactor studied by transient absorption spectroscopy. *ChemPhysChem* **17**, 1329–1340 (2016).

24. Nohr, D. et al. Extended electron-transfer in animal cryptochromes mediated by a tetrad of aromatic amino acids. *Biophys. J.* **111**, 301–311 (2016).

25. Lin, C., Top, D., Manahan, C. C., Young, M. W. & Crane, B. R. Circadian clock activity of cryptochrome relies on tryptophan-mediated

photoreduction. *Proc. Natl Acad. Sci. USA* **115**, 3822–3827 (2018).

26. Gegear, R. J., Foley, L. E., Casselman, A. & Reppert, S. M. Animal cryptochromes mediate magnetoreception by an unconventional photochemical mechanism. *Nature* **463**, 804–807 (2010).

27. Baik, L. S. et al. Distinct mechanisms of *Drosophila* CRYPTOCHROME-mediated light-evoked membrane depolarization and in vivo clock resetting. *Proc. Natl Acad. Sci. USA* **116**, 23339–23344 (2019).

28. Biskup, T. et al. Variable electron transfer pathways in an amphibian cryptochrome tryptophan versus tyrosine-based radical pairs. *J. Biol. Chem.* **288**, 9249–9260 (2013).

29. Solov'yov, I. A. & Schulten, K. Magnetoreception through cryptochrome may involve superoxide. *Biophys. J.* **96**, 4804–4813 (2009).

30. Wiltschko, R., Ahmad, M., Nießner, C., Gehring, D. & Wiltschko, W. Light-dependent magnetoreception in birds: the crucial step occurs in the dark. *J. R. Soc. Interface* **13**, 20151010 (2016).

31. Steiner, U. E. & Ulrich, T. Magnetic field effects in chemical kinetics and related phenomena. *Chem. Rev.* **89**, 51–147 (1989).

32. Woodward, J. R. Radical pairs in solution. *Progress in Reaction Kinetics and Mechanism* **27**, 165–207 (2002).

33. Foley, L. E., Gegear, R. J. & Reppert, S. M. Human cryptochrome exhibits light-dependent magnetosensitivity. *Nat. Commun.* **2**, 356 (2011).

34. Miura, T., Maeda, K. & Arai, T. Effect of Coulomb interaction on the dynamics of the radical pair in the system of flavin mononucleotide and hen egg-white lysozyme (HEWL) studied by a magnetic field effect. *J. Phys. Chem. B* **107**, 6474–6478 (2003).

35. Czarna, A. et al. Structures of *Drosophila* cryptochrome and mouse cryptochrome1 provide insight into circadian function. *Cell* **153**, 1394–1405 (2013).

36. Shi, Z., Olson, C. A. & Kallenbach, N. R. Cation-π interaction in model α-helical peptides. *J. Am. Chem. Soc.* **124**, 3284–3291 (2002).

37. Ikeya, N. & Woodward, J. R. Cellular autofluorescence is magnetic field sensitive. *Proc. Natl Acad. Sci. USA* **118**, e2018043118 (2021).

38. Murakami, M., Maeda, K. & Arai, T. Structure and kinetics of the intermediate biradicals generated from intramolecular electron transfer reaction of FAD studied by an action spectrum of the magnetic field effect. *Chem. Phys. Lett.* **362**, 123–129 (2002).

39. Henbest, K. B., Kukura, P., Rodgers, C. T., Hore, P. J. & Timmel, C. R. Radio frequency magnetic field effects on a radical recombination reaction: a diagnostic test for the radical pair mechanism. *J. Am. Chem. Soc.* **126**, 8102–8103 (2004).

40. Kowalczyk, R. M., Schleicher, E., Bittl, R. & Weber, S. The photoinduced triplet of flavins and its protonation states. *J. Am. Chem. Soc.* **126**, 11393–11399 (2004).

41. Timmel, C. R. et al. Magnetic field effects in flavoproteins and related systems. *Interface Focus* https://doi.org/10.1098/rsfs.2013.0037 (2013).

42. Hemsley, M. J. et al. Linear motifs in the C-terminus of *D. melanogaster* cryptochrome. *Biochem. Biophys. Res. Commun.* **355**, 531–537 (2007).

43. Zito, K., Fetter, R. D., Goodman, C. S. & Isacoff, E. Y. Synaptic clustering of Fasciclin II and Shaker: essential targeting sequences and role of dig. *Neuron* **19**, 1007–1016 (1997).

44. Thomas, U. et al. Synaptic targeting and localization of Discs-large is a stepwise process controlled by different domains of the protein. *Curr. Biol.* **10**, 1108–1117 (2000).

45. Busza, A., Emery-Le, M., Rosbash, M. & Emery, P. Roles of the two *Drosophila* CRYPTOCHROME structural domains in circadian photoreception. *Science* **304**, 1503–1506 (2004).

46. Xu, J. et al. Magnetic sensitivity of cryptochrome 4 from a migratory songbird. *Nature* **594**, 535–540 (2021).

47. Fogle, K. J. et al. Cryptochrome-mediated phototransduction by modulation of the potassium ion channel β-subunit redox sensor. *Proc. Natl Acad. Sci. USA* **112**, 2245–2250 (2015).

48. Pooam, M. et al. Magnetic sensitivity mediated by the *Arabidopsis* blue-light receptor cryptochrome occurs during flavin reoxidation in the dark. *Planta* **249**, 319–332 (2019).

49. Hammad, M. et al. Cryptochrome mediated magnetic sensitivity in *Arabidopsis* occurs independently of light-induced electron transfer to the flavin. *Photochem. Photobiol. Sci.* **19**, 341–352 (2020).

50. Netušil, R. et al. Cryptochrome-dependent magnetoreception in heteropteran insect continues even after 24 hours in darkness. *J. Exp. Biol.* **224**, jeb243000 (2021).

51. Toma, D. P., White, K. P., Hirsch, J. & Greenspan, R. J. Identification of genes involved in *Drosophila melanogaster* geotaxis, a complex behavioral trait. *Nat. Genet.* **31**, 349–353 (2002).

52. Fedele, G., Green, E. W., Rosato, E. & Kyriacou, C. P. An electromagnetic field disrupts negative geotaxis in *Drosophila* via a CRY-dependent pathway. *Nat. Commun.* **5**, 4391 (2014).

53. Bae, J. E. et al. Positive geotactic behaviors induced by geomagnetic field in *Drosophilia*. *Mol. Brain* **9**, 55 (2016).

54. Oh, I. T. et al. Behavioral evidence for geomagnetic imprinting and transgenerational inheritance in fruit flies. *Proc. Natl Acad. Sci. USA* **117**, 1216–1222 (2020).

55. Kyriacou, C. P. & Rosato, E. Genetic analysis of cryptochrome in insect magnetosensitivity. *Front. Physiol.* **13**, 928416 (2022).

56. Dolezelova, E., Dolezel, D. & Hall, J. C. Rhythm defects caused by newly engineered null mutations in *Drosophila's cryptochrome* gene. *Genetics* **177**, 329–345 (2007).

57. Gegear, R. J., Casselman, A., Waddell, S. & Reppert, S. M. Cryptochrome mediates light-dependent magnetosensitivity in *Drosophila*. *Nature* **454**, 1014–1018 (2008).

58. Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).

59. Baines, R. A. & Bate, M. Electrophysiological development of central neurons in the *Drosophila* embryo. *J. Neurosci.* **18**, 4673–4683 (1998).

60. Delfino, L. et al. Visualization of mutant aggregates from clock neurons by agarose gel electrophoresis (AGERA) in *Drosophila melanogaster*. *Methods Mol. Biol.* **2482**, 373–383 (2022).

Acknowledgements

A.R.J. thanks I. Camacho and M. Shaw for feedback on the manuscript. This work was supported by funding from the Leverhulme Trust to R.A.B. and A.R.J. (RPG-2017-113), to R.A.B. from BBSRC (BB/V005987/1), and to C.P.K. and E.R. from BBSRC (BB/V006304/1). A.L.M. and J.M.H. were funded by BBSRC studentships. Work on this project benefited from the Manchester Fly Facility, established through funds from the University and the Wellcome Trust (087742/Z/08/Z). A.R.J. acknowledges the National Measurement System of the Department for Business, Energy and Industrial Strategy for funding. G.F., C.N.H., E.R. and C.P.K. acknowledge funding from the Electromagnetic Field Biological Research Trust (BRT 15/51).

Author information

Authors and Affiliations

1. Division of Neuroscience, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK

   Adam A. Bradlaugh, Anna L. Munro & Richard A. Baines

2. Department of Genetics, University of Leicester, Leicester, UK

   Giorgio Fedele, Celia Napier Hansen, John M. Hares, Charalambos P. Kyriacou & Ezio Rosato

3. Department of Biology, Ecology and Earth Sciences (DiBEST), University of Calabria, Rende, Italy

   Giorgio Fedele

4. Pelican Healthcare, Cardiff, UK

   John M. Hares

5. Manchester Fly Facility, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

   Sanjai Patel

6. Biometrology, Chemical and Biological Sciences Department, National Physical Laboratory, Teddington, UK

   Alex R. Jones

Contributions

G.F., A.L.M. and C.N.H. performed experiments. S.P. and J.M.H. generated transgenic *Drosophila*. A.A.B., C.P.K., G.F., E.R. and R.A.B analysed the data. A.A.B., C.P.K., A.R.J., E.R. and R.A.B. wrote the manuscript. C.P.K., E.R., A.R.J. and RAB acquired funding.

Corresponding authors

Correspondence to Ezio Rosato or Richard A. Baines.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks the anonymous reviewers for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Supporting electrophysiological data for main text Fig. 1.

(a). Raw action potential (AP) counts for aCC neurons expressing *Dm*CRY, BL: n = 10, BL+MF: n = 10 (b). Luc-CT, BL: n = 10, BL+MF: n = 10 or (c). Luc-CT^{W536F}, BL: n = 20, BL+MF: n = 20, recorded in the 15 s before and 15 s during exposure to BL or BL±MF. These data were used to derive firing-fold (FF) change reported in main text Fig. 1. Paired t-tests – two tailed – were used to compare AP counts before *vs*. during for cells exposed to BL (left hand graph) or to BL±MF exposure (right hand graph). MF-potentiation between the two groups was tested by unpaired t-test – two

tailed (different cells). The figure number at the top of each panel represents the main text figure supported. The reported n for each electrophysiological recording is derived from independent cells from biologically independent animals. ns $p \geq 0.06$, *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. See main text Fig. 1 for FF_{on}/FF_{off} ratio comparisons.

Extended Data Fig. 2 Supporting circadian data for main text Fig. 1.

(a). Showing a relative period shortening following expression of Luc-CT under MF (EMF- bars coloured red) compared to the BL+Sham exposure (bars coloured blue). (b). Box and whisker diagram (95% confidence limits) representation of the data shown in the upper panel, omitting DD. Error bars denote ±SEM, ns $p \geq 0.06$, *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$.

Extended Data Fig. 3 Supporting electrophysiological data (controls).

(a). Averaged data for parental (control) *GAL4* and *UAS* genotypes separately in a *Dmcry*-null mutant background. Without the *GAL4* and *UAS* elements combined the *Dmcry*-transgene (under *UAS* control) is not expressed and no significant BL or BL+MF response is seen. A 2-way ANOVA of the controls in both BL or BL+MF showed no significant interaction ($F_{(4,50)} = 0.52$, $p = 0.719$), n for all UAS-*Dmcry* transgenic controls in BL and BL+MF = 5. n for *elav-GAL4; ; Dmcry^{02}/Dmcry^{03}* driver line BL and BL+MF = 10. (b-c). Raw AP counts *(i)* for each aCC neuron recorded in the 15 s before *vs*. the 15 s during BL or BL±MF exposure for expression of the respective UAS-*Dmcry* transgene stated (paired t-tests – two tailed). As an additional comparison of MF effect, unpaired t-tests – two tailed, were used to determine significant differences between raw AP counts for both ‘before’ exposure conditions (grey line), and for during BL *vs*. BL+MF exposure (purple line). The reported n for each electrophysiological recording is derived from independent cells from biologically independent animals. The reported n for each circadian period derives from biologically independent animals. Error bars denote ±SEM, ns $p \geq 0.06$, *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$.

Extended Data Fig. 4 Supporting electrophysiological data (controls).

(a–c). Raw AP counts *(i)* for each aCC neuron recorded in the 15 s before *vs*. the 15 s during BL or BL±MF exposure for expression of the respective UAS-*Dmcry* transgene stated (paired t-tests – two tailed). (d). Data for a *Dmcry* null. As an additional comparison of MF effect, unpaired t-tests – two tailed were used to determine significant differences between raw AP counts for both 'before' exposure conditions (grey line), and for during BL *vs*. BL+MF exposure (purple line). The reported n for each electrophysiological recording is derived from independent cells from biologically independent animals. The reported n for each circadian period derives from biologically independent animals. Error bars denote ±SEM, ns $p \geq 0.06$, $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$.

Extended Data Fig. 5 Supporting electrophysiological data for main text Fig. 2.

Raw AP counts for averaged data shown for flavin supplementation. (a). *Dmcry^{02}/Dmcry^{03}* null cells supplemented with 50 μM FAD show a BL response ($n = 10$), but no MF potentiation ($n = 10$) compared to BL alone. (b). Cells expressing Luc-CT supplemented with riboflavin (50 μM) show no MF effect ($n = 10$) compared to BL ($n = 10$). (c). Supplemented FAD and riboflavin (200 μM) to *Dmcry^{02}/Dmcry^{03}* null cells: FAD supports BL ($n = 20$) and BL+MF ($n = 20$) sensitivity, whilst riboflavin only supports BL sensitivity (see main text for BL and BL+MF comparisons). Paired t-tests – two tailed, were used to compare before *vs*. during for cells exposed to BL (left hand graph) or to BL±MF exposure (right hand graph). MF-potentiation between the two groups was tested by unpaired t-tests – two tailed (different cells). The reported n for each electrophysiological recording is derived from independent cells from biologically independent animals. ns $p \geq 0.06$, $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$.

Extended Data Fig. 6 Structures of FAD and Riboflavin.

The molecular structure of the two flavin chromophores used. (a). Flavin Adenine Dinucleotide (FAD). Note the adenine diphosphate side chain of FAD (yellow oval), which facilitates the generation of an intramolecular, magnetically-sensitive RP. (b). Riboflavin, which is a metabolic precursor to FAD lacks the diphosphate side chain.

Extended Data Fig. 7 Supporting circadian and electrophysiology data for main text Figs. 3 and 4.

(a). Raw AP counts for each aCC neuron recorded expressing $Dmcry^{V531K}$ in the 15 s before *vs*. the following 15 s during BL±MF, from which the average firing-fold change was derived for data reported in main text Fig. 3c. BL: n = 10, BL+MF: n = 10. Paired t-tests – two tailed, were used to compare before *vs*. during for cells exposed to BL (left hand graph) or to BL±MF exposure (right hand graph). MF-potentiation between the two groups was tested by unpaired t-tests – two tailed (different cells). (b). *Dm*cry^{M} flies do not undergo period shortening following exposure to a 50 µT MF (DD n = 26, BL Pre-Sham n = 51, BL+Sham n = 42, BL Pre-MF n = 53, BL+MF n = 38). (c). Raw AP counts for BL and BL+MF exposure for *Er*CRY4, showing an increase in neuronal excitability to both stimuli (BL: n = 8, BL+MF: n = 9). Paired t-tests – two tailed were used to compare before *vs*. during for cells exposed to BL (left hand graph) or to BL±MF exposure (right hand graph). MF-potentiation between the two groups was tested by unpaired t-tests – two tailed (different cells). The reported n for each electrophysiological recording is derived from independent cells from biologically independent animals. The reported n for each circadian period derives from biologically independent animals. Error bars denote ±SEM, ns $p \geq 0.06$, *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$.

Extended Data Table 1 Supporting circadian data for main text Fig. 1
Extended Data Table 2 Supporting circadian data for main text Fig. 1
Extended Data Table 3 Supporting electrophysiological controls
Extended Data Table 4 Supporting circadian data for main text Fig. 3
Extended Data Table 5 Circadian and data for *Er*CRY4 in main text Fig. 4

Supplementary information

[Reporting Summary](#)

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Bradlaugh, A.A., Fedele, G., Munro, A.L. *et al.* Essential elements of radical pair magnetosensitivity in *Drosophila*. *Nature* **615**, 111–116 (2023). https://doi.org/10.1038/s41586-023-05735-z

- Received: 28 October 2021
- Accepted: 16 January 2023
- Published: 22 February 2023
- Issue Date: 02 March 2023
- DOI: https://doi.org/10.1038/s41586-023-05735-z

Share this article

Anyone you share the following link with will be able to read this content:

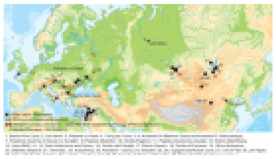
Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

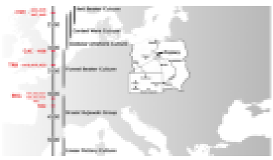
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05735-z



Initial Upper Palaeolithic humans in Europe had recent Neanderthal ancestry

07 April 2021

Mateja Hajdinjak, Fabrizio Mafessoni, … Svante Pääbo



A genomic Neolithic time transect of hunter-farmer admixture in central Poland

05 October 2018

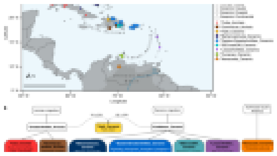
D. M. Fernandes, D. Strapagiel, … R. Pinhasi



Palaeo-Eskimo genetic ancestry and the peopling of Chukotka and North America

05 June 2019

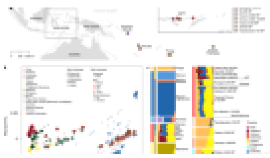
Pavel Flegontov, N. Ezgi Altınışık, … Stephan Schiffels



[A genetic history of the pre-contact Caribbean]

23 December 2020

Daniel M. Fernandes, Kendra A. Sirak, … David Reich



[Ancient genomes from the last three millennia support multiple human dispersals into Wallacea]

09 June 2022

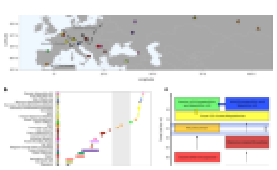
Sandra Oliveira, Kathrin Nägele, … Mark Stoneking



[Ancient DNA reveals admixture history and endogamy in the prehistoric Aegean]

16 January 2023

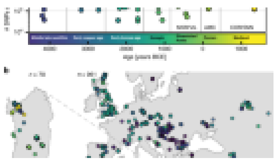
Eirini Skourtanioti, Harald Ringbauer, … Philipp W. Stockhammer



[A 23,000-year-old southern Iberian individual links human groups that lived in Western Europe before and after the Last Glacial Maximum]

01 March 2023

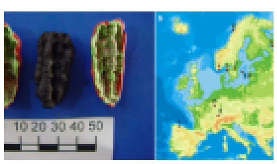
Vanessa Villalba-Mouco, Marieke S. van de Loosdrecht, … Wolfgang Haak



[Genetic history from the Middle Neolithic to present on the Mediterranean island of Sardinia]

24 February 2020

Joseph H. Marcus, Cosimo Posth, … John Novembre



[Ancient DNA from mastics solidifies connection between material culture and genetics of mesolithic hunter–gatherers in Scandinavia]

15 May 2019

Natalija Kashuba, Emrah Kırdök, … Anders Götherström

- Article
- Open Access
- [Published: 01 March 2023]

Palaeogenomics of Upper Palaeolithic to Neolithic European hunter-gatherers

- Cosimo Posth ,
- He Yu ,
- Ayshin Ghalichi,
- Hélène Rougier,
- Isabelle Crevecoeur,
- Yilei Huang,
- Harald Ringbauer,
- Adam B. Rohrlach,
- Kathrin Nägele,
- Vanessa Villalba-Mouco,
- Rita Radzeviciute,
- Tiago Ferraz,
- Alexander Stoessel,
- Rezeda Tukhbatova,
- Dorothée G. Drucker,
- Martina Lari,
- Alessandra Modi,
- Stefania Vai,
- Tina Saupe,
- Christiana L. Scheib,
- Giulio Catalano,
- Luca Pagani,
- Sahra Talamo,
- Helen Fewlass,
- Laurent Klaric,
- André Morala,
- Mathieu Rué,
- Stéphane Madelaine,
- Laurent Crépin,
- Jean-Baptiste Caverne,

- Emmy Bocaege,
- Stefano Ricci,
- Francesco Boschin,
- Priscilla Bayle,
- Bruno Maureille,
- Foni Le Brun-Ricalens,
- Jean-Guillaume Bordes,
- Gregorio Oxilia,
- Eugenio Bortolini,
- Olivier Bignon-Lau,
- Grégory Debout,
- Michel Orliac,
- Antoine Zazzo,
- Vitale Sparacello,
- Elisabetta Starnini,
- Luca Sineo,
- Johannes van der Plicht,
- Laure Pecqueur,
- Gildas Merceron,
- Géraldine Garcia,
- Jean-Michel Leuvrey,
- Coralie Bay Garcia,
- Asier Gómez-Olivencia,
- Marta Połtowicz-Bobak,
- Dariusz Bobak,
- Mona Le Luyer,
- Paul Storm,
- Claudia Hoffmann,
- Jacek Kabaciński,
- Tatiana Filimonova,
- Svetlana Shnaider,
- Natalia Berezina,
- Borja González-Rabanal,
- Manuel R. González Morales,
- Ana B. Marín-Arroyo,
- Belén López,
- Carmen Alonso-Llamazares,

- Annamaria Ronchitelli,
- Caroline Polet,
- Ivan Jadin,
- Nicolas Cauwe,
- Joaquim Soler,
- Neus Coromina,
- Isaac Rufí,
- Richard Cottiaux,
- Geoffrey Clark,
- Lawrence G. Straus,
- Marie-Anne Julien,
- Silvia Renhart,
- Dorothea Talaa,
- Stefano Benazzi,
- Matteo Romandini,
- Luc Amkreutz,
- Hervé Bocherens,
- Christoph Wißing,
- Sébastien Villotte,
- Javier Fernández-López de Pablo,
- Magdalena Gómez-Puche,
- Marco Aurelio Esquembre-Bebia,
- Pierre Bodu,
- Liesbeth Smits,
- Bénédicte Souffi,
- Rimantas Jankauskas,
- Justina Kozakaitė,
- Christophe Cupillard,
- Hartmut Benthien,
- Kurt Wehrberger,
- Ralf W. Schmitz,
- Susanne C. Feine,
- Tim Schüler,
- Corinne Thevenet,
- Dan Grigorescu,
- Friedrich Lüth,
- Andreas Kotula,

- Henny Piezonka,
- Franz Schopper,
- Jiří Svoboda,
- Sandra Sázelová,
- Andrey Chizhevsky,
- Aleksandr Khokhlov,
- Nicholas J. Conard,
- Frédérique Valentin,
- Katerina Harvati,
- Patrick Semal,
- Bettina Jungklaus,
- Alexander Suvorov,
- Rick Schulting,
- Vyacheslav Moiseyev,
- Kristiina Mannermaa,
- Alexandra Buzhilova,
- Thomas Terberger,
- David Caramelli,
- Eveline Altena,
- Wolfgang Haak &
- …
- Johannes Krause

Show authors

Nature **volume 615**, pages 117–126 (2023)

- 24k Accesses

- 1075 Altmetric

- Metrics details

Abstract

Modern humans have populated Europe for more than 45,000 years[1,2]. Our knowledge of the genetic relatedness and structure of ancient hunter-

gatherers is however limited, owing to the scarceness and poor molecular preservation of human remains from that period[3]. Here we analyse 356 ancient hunter-gatherer genomes, including new genomic data for 116 individuals from 14 countries in western and central Eurasia, spanning between 35,000 and 5,000 years ago. We identify a genetic ancestry profile in individuals associated with Upper Palaeolithic Gravettian assemblages from western Europe that is distinct from contemporaneous groups related to this archaeological culture in central and southern Europe[4], but resembles that of preceding individuals associated with the Aurignacian culture. This ancestry profile survived during the Last Glacial Maximum (25,000 to 19,000 years ago) in human populations from southwestern Europe associated with the Solutrean culture, and with the following Magdalenian culture that re-expanded northeastward after the Last Glacial Maximum. Conversely, we reveal a genetic turnover in southern Europe suggesting a local replacement of human groups around the time of the Last Glacial Maximum, accompanied by a north-to-south dispersal of populations associated with the Epigravettian culture. From at least 14,000 years ago, an ancestry related to this culture spread from the south across the rest of Europe, largely replacing the Magdalenian-associated gene pool. After a period of limited admixture that spanned the beginning of the Mesolithic, we find genetic interactions between western and eastern European hunter-gatherers, who were also characterized by marked differences in phenotypically relevant variants.

Main

Modern humans left sub-Saharan Africa at least 60 thousand years ago (ka), and during their initial expansion into Eurasia, they genetically mixed with Neanderthals, resulting in 2–3% Neanderthal ancestry in the majority of present-day non-African populations[5]. Genomic data have shown that modern humans were present in western Eurasia[1,2] at least 45 ka. Some of those early groups from more than 40 ka further admixed with Neanderthals, as shown by signals of recent introgression in individuals from Bacho Kiro in Bulgaria—associated with an Initial Upper Palaeolithic (IUP) archaeological culture—and from Peştera cu Oase in Romania[2,6]. Other individuals from that period, such as Zlatý kůň from Czechia and Ust'Ishim

from Russia, do not carry significantly more Neanderthal ancestry than other non-African groups[1,7], indicating differential interactions between Neanderthals and early modern humans during their initial expansions across Eurasia. Surprisingly, however, none of those pre-40 ka individuals left substantial traces in the genetic makeup of present-day Eurasian populations[1,2,6,7]. The oldest genomes carrying ancestries that derive primarily from the lineage leading to present-day Europeans are Kostenki 14 (from 37 ka, with uncertain archaeological association from western Russia), Goyet Q116-1 (35 ka, Aurignacian-associated from Belgium) and Bacho Kiro 1653 (35 ka, probably Aurignacian-associated from Bulgaria)[2,4,8]. These data suggest that the genetic ancestries identified in the pre-40 ka individuals analysed so far went largely extinct or were assimilated by subsequent expansions[1,9]. The Kostenki genetic signature (related to the Kostenki 14 genome, and hereafter referred to as the Kostenki cluster or ancestry) contributed to the later Věstonice genetic cluster (hereafter, Věstonice cluster or ancestry), named after the Dolní Věstonice site in Czechia[4]. This genetic signature is shared among individuals associated with the archaeologically defined Gravettian culture (33–26 ka) in central and southern Europe and seemingly disappeared after the Last Glacial Maximum[4] (LGM). However, the genetic profile of contemporaneous Gravettian-associated individuals from western Europe remains unknown, as is their contribution to populations after the LGM. Known to have been the coldest phase of the last Ice Age, the LGM is considered to have caused a demographic decline in large parts of Europe[10], with populations retracting to southern latitudes as attested—for example—by the contemporaneity of the Solutrean culture (24–19 ka) in the Iberian peninsula and southern France. Other proposed climatic refugia for human survival during this period are the Italian peninsula, the Balkans and the southeastern European Plain, but the actual genetic contribution of populations from these regions to post-LGM Europeans is highly debated[11,12,13].

After the LGM, a genetic component distantly linked to the Goyet Q116-1 individual from Belgium dated to 35 ka—named GoyetQ2 ancestry (hereafter, GoyetQ2 cluster or ancestry)—reappeared in individuals from southwestern and central Europe associated with the Magdalenian culture (19–14 ka from Iberia to eastern Europe across central Europe) and

in an admixed form in subsequent Final Palaeolithic and Mesolithic hunter-gatherers[4,14], but the geographic extension of this ancestry is still unclear. Instead, in southern Europe, a distinct hunter-gatherer genetic profile was found as early as 17 ka in individuals associated with the Epigravettian culture[15] (24–12 ka, from the Italian peninsula to the southeastern European Plain across the Balkans). This 'Villabruna' ancestry (hereafter, Villabruna cluster or ancestry) showed connections to ancient and present-day Near Eastern populations[4,16], but the mode and tempo of its expansion into the Italian peninsula remain unexplored. The Villabruna ancestry later appeared in central Europe and it is thought to have largely replaced groups related to the GoyetQ2 ancestry[4]. However, its formation, diffusion and interaction with contemporaneous hunter-gatherers from eastern Europe and their interplay with later expansions of Neolithic farmers from southeastern Europe are not well characterized.

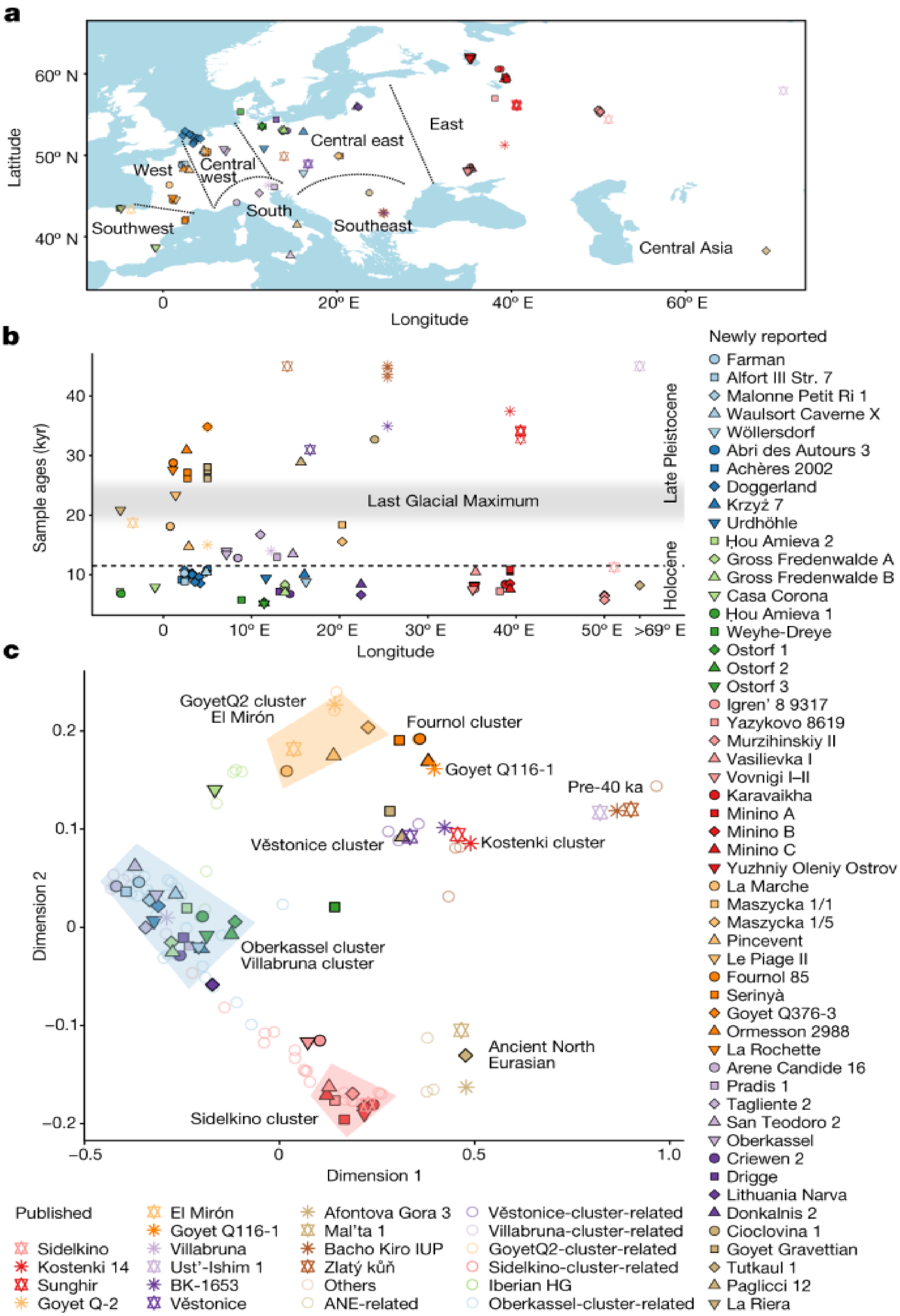
In this study, we analyse 356 ancient hunter-gatherer genomes including new genomic data of 116 individuals dated to 35–5 ka alongside a novel contamination-estimation method based on runs of homozygosity. We provide a systematic description of the genomic transformations that hunter-gatherer groups experienced from the early Upper Palaeolithic onwards across western and central Eurasia and how those are possibly linked to cultural and climatic changes.

Ancient DNA data generation

We generated genome-wide sequencing data for 102 newly reported hunter-gatherers, and increased coverage for 14 previously published individuals[4]. These data cover a time span of around 30,000 years from the Upper Palaeolithic to the Late Neolithic (defined here by the presence of pottery rather than by farming subsistence economy if not indicated), derive from multiple prehistoric cultural contexts, and originate from 54 archaeological sites in 14 countries: 1 Aurignacian-associated individual from Belgium and 1 culturally unassigned individual from Romania (35–33 ka), 15 Gravettian-associated individuals from Spain, France, Belgium, Czechia and Italy (31–26 ka), 2 Solutrean-associated individuals from Spain and France (23–21 ka), 9 Magdalenian-associated individuals from France, Germany, and Poland (18–15 ka), 4 Epigravettian-associated individuals from Italy (17–13

ka), 2 Federmesser-associated individuals from Germany (14 ka), and 81 Mesolithic to Neolithic foragers from across western Eurasia (11–5 ka), together with 1 central Eurasian Neolithic individual from Tajikistan (8 ka) (Fig. 1, Extended Data Table 1, Supplementary Data 1.A, Supplementary Information, section 1 and Supplementary Fig. 1).

Fig. 1: Locations, dates and MDS plot of ancient Eurasian hunter-gatherers.



a
Latitude
60° N
50° N
40° N
0
20° E
40° E
60° E
Longitude
West
Central west
Central east
East
South
Southwest
Southeast
Central Asia
b
Sample ages (kyr)
Last Glacial Maximum
Late Pleistocene
Holocene
10° E
30° E
50° E
>69° E
c
Dimension 2
Dimension 1
GoyetQ2 cluster
El Mirón
Fournol cluster
Goyet Q116-1
Pre-40 ka
Věstonice cluster
Kostenki cluster
Oberkassel cluster
Villabruna cluster
Ancient North Eurasian
Sidelkino cluster
Newly reported
Farman
Alfort III Str. 7
Malonne Petit Ri 1
Waulsort Caverne X
Wöllersdorf
Abri des Autours 3
Achères 2002
Doggerland
Krzyż 7
Urdhöhle
Ḥou Amieva 2
Gross Fredenwalde A
Gross Fredenwalde B
Casa Corona
Ḥou Amieva 1
Weyhe-Dreye
Ostorf 1
Ostorf 2
Ostorf 3
Igren' 8 9317
Yazykovo 8619
Murzihinskiy II
Vasilievka I
Vovnigi I–II
Karavaikha
Minino A
Minino B
Minino C
Yuzhniy Oleniy Ostrov
La Marche
Maszycka 1/1
Maszycka 1/5
Pincevent
Le Piage II
Fournol 85
Serinyà
Goyet Q376-3
Ormesson 2988
La Rochette
Arene Candide 16
Pradis 1
Tagliente 2
San Teodoro 2
Oberkassel
Criewen 2
Drigge
Lithuania Narva
Donkalnis 2
Cioclovina 1
Goyet Gravettian
Tutkaul 1
Paglicci 12
La Riera
Published
Sidelkino
Kostenki 14
Sunghir
Goyet Q-2
El Mirón
Goyet Q116-1
Villabruna
Ust'-Ishim 1
BK-1653
Věstonice
Afontova Gora 3
Mal'ta 1
Bacho Kiro IUP
Zlatý kůň
Others
ANE-related
Věstonice-cluster-related
Villabruna-cluster-related
GoyetQ2-cluster-related
Sidelkino-cluster-related
Iberian HG
Oberkassel-cluster-related

a, Geographic locations of newly reported individuals (filled symbols with black outline) and representative previously published individuals (outlined stars). Dotted lines delimit geographic regions described in the text. **b**, Calibrated radiocarbon dates of individuals plotted in **a**. The *y* axis shows the average of calibrated radiocarbon dates in thousands of years (kyr) (Supplementary Data 1.A). The horizontal dashed line marks the boundary between Late Pleistocene and Holocene. **c**, MDS plot of European hunter-gatherers based on $1 - f_3$(Mbuti; pop1, pop2). The dimensions are calculated using newly reported and previously published hunter-gatherer groups or individuals with more than 30,000 SNPs. The detailed grouping of individuals shown with empty coloured circles is described in Supplementary Data 1.I.

We built 1 to 8 single- and double-stranded genetic libraries for each individual and enriched them for human DNA on 1.24 million single nucleotide polymorphisms[6] (SNPs), which were then sequenced and yielded 0.04- to 7.64-fold coverage on average over the targeted SNPs. Genetic sexing revealed 78 male individuals and 38 female individuals (Supplementary Fig. 12). The levels of contamination from modern human DNA were estimated on the basis of mitochondrial DNA (mtDNA), X chromosome and autosomal DNA, and with a haplotype copying model that is extended here to autosomal data in runs of homozygosity (ROH) (Methods, Supplementary Information, sections 2 and 3, Supplementary Figs. 2–11 and Supplementary Table 1). Substantially contaminated libraries as well as marginally contaminated libraries of individually analysed genomes were filtered to maintain reads showing postmortem DNA damage (Methods and Supplementary Figs. 10 and 11). Pseudo-haploid genotypes were called on the targeted SNPs by randomly sampling a single allele at each position, resulting in individuals with 6,600 to 1.07 million SNPs covered on the 1.24-million-SNP panel (Extended Data Table 1 and Supplementary Data 1.A). The newly generated genotypes were merged with 240 published ancient hunter-gatherer genomes and modern worldwide populations for downstream analyses (Supplementary Data 1.G). Contrary to the proposal in Fu et al.[4] but in agreement with Petr et al.[17], we do not observe a substantial decrease of Neanderthal ancestry in most European hunter-gatherers through time (Supplementary Information, section 6 and Supplementary Figs. 15–17). This provides further support for the model

with no long-term decline of genome-wide Neanderthal ancestry in modern humans following their introgression[18].

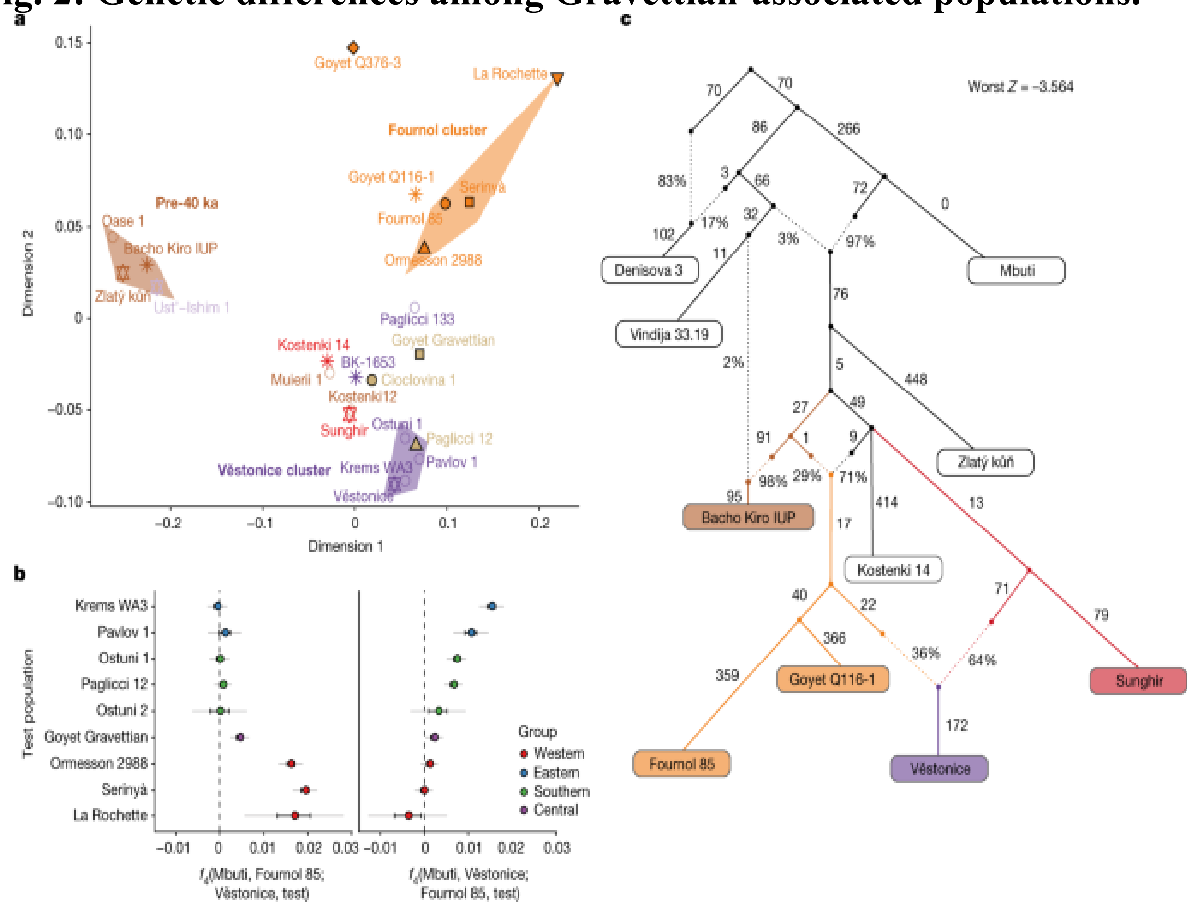
Before the LGM

The Gravettian culture was one of the most widely distributed Upper Palaeolithic cultures across western Eurasia before the LGM[19]. It is often considered as a pan-European cultural mosaic with regional variations in material to symbolic productions[20,21]. In this debated framework, Gravettian-associated individuals have been suggested to represent a biologically homogeneous population on the basis of craniometric and genomic data[4,22]. However, published Gravettian-associated genomes originate from central and southern Europe, leaving the genetic profile of Gravettian-associated human groups from western and southwestern Europe undescribed.

To gain an overview of the genomic background of European hunter-gatherers before the LGM, we used multidimensional scaling (MDS) to plot a dissimilarity matrix of pairwise outgroup f_3-statistics in the form $1 - f_3$(Mbuti; pop1, pop2) (Fig. 2a). This plot reveals the presence of three distinct groupings: (1) a pre-40 ka group with individuals from the Ust'Ishim, Bacho Kiro, Zlatý kůň and Peştera cu Oase sites, (2) a Věstonice cluster including Gravettian-associated individuals from central–eastern and southern European sites (Dolní Věstonice, Pavlov, Krems-Wachtberg, Paglicci and Ostuni), and (3) a Fournol cluster (hereafter, Fournol cluster or ancestry) comprising Gravettian-associated individuals from western and southwestern European sites (Ormesson, La Rochette, Fournol and two Serinyà cave sites (Mollet III and Reclau Viver)). The previously described Věstonice cluster, including a newly reported 29,000-year-old individual from Paglicci cave (Paglicci 12) in southern Italy, is closely related to the previously published genomes from Sunghir and Kostenki 12 in western Russia, which are dated to 34 ka and 32 ka, respectively[4,23]. The newly defined Fournol cluster is closely related to Aurignacian-associated individuals from Belgium dated to 35 ka (Goyet Q116-1 and the newly reported Goyet Q376-3 individual). Notably, and contrary to the report by Fu et al.[4], another Gravettian-associated population from central–western

Europe (Goyet in Belgium, $n = 6$ individuals) is both geographically and genetically intermediate between the Věstonice and Fournol clusters. The similarity between Goyet Q116-1 and Goyet Q376-3 and the Fournol cluster is also observed at the mtDNA level, with both groups including individuals who carried mtDNA haplogroup M, which has not been found in European individuals from after the LGM[24] (Extended Data Figs. 1 and 2).

Fig. 2: Genetic differences among Gravettian-associated populations.



a, MDS plot of pre-LGM individuals. The pre-40 ka group and the Fournol and Věstonice clusters are marked as shaded areas in different colours. Individuals and groups are plotted with the same colours and symbols as in Fig. 1 and names are indicated next to the symbols. **b**, Gravettian-associated individuals form two distinct groups, with central-eastern and southern European individuals as part of the Věstonice cluster and western and southwestern European individuals as part of the Fournol cluster. In central-western Europe, Gravettian-associated individuals from Goyet show affinity to both clusters. Error bars show 1× s.e.m. (black) or 3× s.e.m. (grey) of the

f_4 values estimated from 5 cM-block jackknife analysis. **c**, Admixture graph modelling of the main pre-LGM European hunter-gatherer lineages created using qpGraph.

We further validated the genetic distinction between the Věstonice and Fournol clusters observed in the MDS plot with a series of f_4-statistics (Supplementary Data 2.B). All individuals belonging to the Fournol cluster show higher affinity to Goyet Q116-1 than to the Sunghir group ($n = 4$), and the Věstonice-cluster individuals show higher affinity to the Sunghir group than to Goyet Q116-1 (Extended Data Fig. 3). These f_4-statistics also confirm that Goyet Q376-3 carries a similar ancestry to Goyet Q116-1 and Kostenki 12 carries a similar ancestry to the Sunghir group, whereas Bacho Kiro 1653 (35 ka) from Bulgaria, Muierii 1 (34 ka) and Cioclovina 1 (32 ka) from Romania, and Paglicci 133 (33 ka) from southern Italy are equally related to Goyet Q116-1 and Sunghir. We further tested whether individuals included in the Věstonice and Fournol clusters share similar allele frequencies with the main representatives of those two clusters. With the statistics f_4(Mbuti, Fournol 85; Věstonice, test) and f_4(Mbuti, Věstonice; Fournol 85, test), we show that all Věstonice-cluster individuals are significantly closer ($|Z|>3$) to the Věstonice group ($n = 5$) and the Fournol-cluster individuals are closer to Fournol 85, whereas the geographically intermediate Gravettian-associated Goyet group shows extra affinity to both clusters (Fig. 2b).

We further modelled the genetic profile of pre-LGM individuals with qpGraph (Supplementary Information, section 10 and Supplementary Figs. 19–25). The admixture graph shows that the Bacho Kiro IUP group ($n = 3$) shares ancestry with multiple early modern human lineages[2] (Supplementary Information, section 7), and that the more than 45,000-year-old Zlatý kůň genome[1] is the most deeply divergent non-African lineage sequenced to date (Extended Data Fig. 4). This is also validated by f_4-statistics of the form f_4(Mbuti, Zlatý kůň; test1, test2), which are consistent with zero for all other pre-LGM hunter-gatherers (Supplementary Data 2.C), indicating an equidistant relationship of Zlatý kůň to the tested groups. When Gravettian-associated individuals are included in an admixture graph also featuring Kostenki 14, we find that Fournol 85 fits best as a sister lineage of Goyet

Q116-1, whereas the Věstonice group is modelled as a mixture between a lineage related to the Sunghir group and one related to the Goyet Q116-1–Fournol 85 branch (Fig. 2c). This is also supported by f_4-statistics of the form f_4(Mbuti, Fournol 85; Sunghir, test), which are significantly positive for all the individuals included in the Věstonice cluster (Supplementary Data 2.B). Therefore, as previously reported[2], the Věstonice cluster itself results from admixture between western and eastern lineages, which might contribute to the observed homogeneity in cranial morphology among Gravettian-associated individuals[22].

These results show that some, but not all, of the genomic ancestries present in Europe between around 40 ka and 30 ka survived in the Gravettian-associated populations studied so far. The Kostenki (and Sunghir group) ancestry contributed to the previously described Věstonice cluster represented by Gravettian-associated individuals from central-eastern and southern Europe[4]. By contrast, the Goyet Q116-1 genetic profile gave rise to the newly described Fournol cluster identified in Gravettian-associated individuals from western and southwestern Europe. Notably, this genetic distinction coincides with dissimilarities in mortuary practice among genetically analysed Gravettian-associated individuals from different parts of Europe. Individuals in western and southwestern Europe related to the Fournol cluster are consistently deposited in cave sites and occasionally exhibit anthropogenic marks whereas individuals related to the Věstonice cluster are buried with grave goods and/or personal ornaments and ochre in open air or cave sites in central-eastern and southern Europe, respectively (Supplementary Figs. 29–32 and Supplementary Table 4). The oldest individual in the Fournol cluster is Ormesson 2988 from northeastern France (31 ka, Early/Middle Gravettian), whereas a Gravettian group from Goyet in Belgium (27 ka, Late Gravettian) is found to be a mixture between the Věstonice and Fournol clusters. This suggests that between the Early/Middle and Late Gravettian there was an east-to-west expansion of the Věstonice-associated ancestry that reached central-western Europe and created a longitudinal admixture cline between those two genetically distinct pre-LGM populations.

LGM in southwestern and western Europe

The Solutrean culture is temporally intermediate between the Gravettian and the Magdalenian (or the Badegoulian) cultures, and is found in southwestern and western Europe, which are considered to have been climatic refugia for human populations during the LGM[25,26]. However, the extent to which groups associated with the Solutrean culture are in genetic continuity with earlier and later populations from the same region is unknown because no genomic data from Solutrean-associated individuals have been reported previously. Both newly sequenced genomes from Solutrean-associated individuals (Le Piage II (23 ka) from southwestern France and La Riera (level 14, 21 ka) from northern Spain) show a generalized affinity with members of the Fournol and GoyetQ2 clusters in outgroup f_3-statistics (Supplementary Data 2.A). In the MDS plot, the Le Piage II individual falls particularly close to individuals belonging to the Fournol cluster, suggesting a local genetic continuity of this ancestry during the LGM (Supplementary Fig. 13). F_4-statistics further support this view, revealing that Le Piage II is more closely related to the Fournol cluster than the Věstonice cluster (f_4(Mbuti, Le Piage II; Věstonice, Fournol 85) $\gg 0$, $Z = 6.58$). We also compared its affinity to El Mirón (northern Spain), the oldest Magdalenian-associated individual sequenced to date (19 ka). F-statistics suggest that Le Piage II is genetically intermediate between Fournol 85 and El Mirón (Supplementary Data 2.D). Moreover, previous studies have shown that El Mirón carries a genetic contribution from the Villabruna cluster, which is found in Epigravettian-associated individuals from Italy[4,15]. El Mirón has a significantly higher similarity to the Villabruna cluster than Fournol 85 and Le Piage II, while the affinity to the Villabruna cluster in Le Piage II is not significantly higher than in Fournol 85 (Supplementary Data 2.D). Overall, the Solutrean-associated Le Piage II individual links the preceding Fournol ancestry with the succeeding ancestry found in El Mirón, providing direct evidence for genetic continuity throughout the LGM in southwestern and western Europe. These European regions, therefore, constitute climatic refugia where human populations survived during the LGM.

Post-LGM in the Italian peninsula

After the LGM, the Epigravettian culture was widespread in southern and southeastern Europe. In spite of growing discussions about its

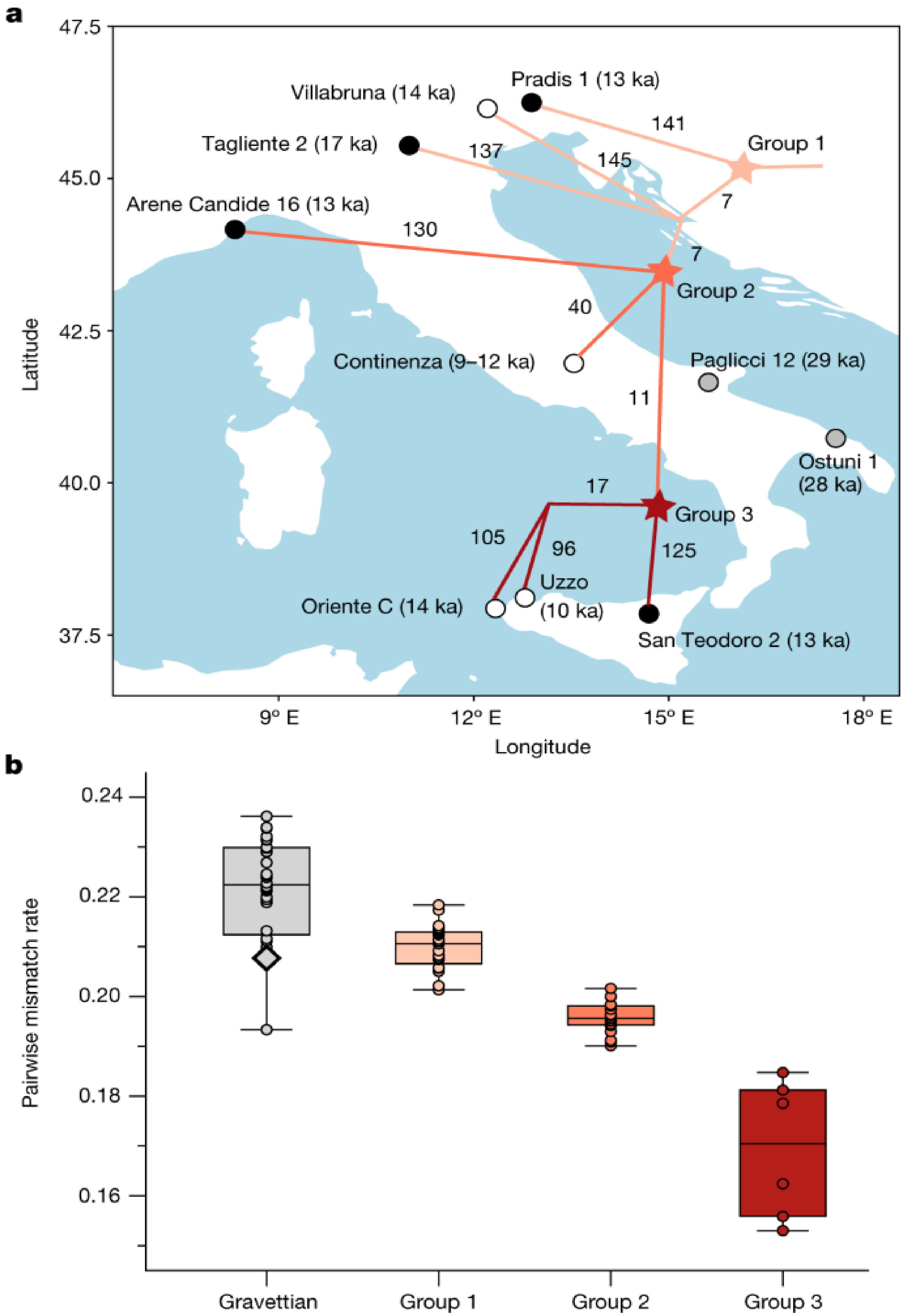
nature[27,28], the Epigravettian culture has been traditionally assumed to be the result of a transition from the preceding local Gravettian[29]. However, the level of genetic continuity between individuals associated with these cultures and the population structure among Epigravettian-associated individuals have not been fully explored. Here, we report genomic data from 4 individuals, including 3 approximately 13,000-year-old genomes from northeastern Italy (Pradis 1), northwestern Italy (Arene Candide 16) and Sicily (San Teodoro 2), as well as increased genome-wide coverage from Tagliente 2[15] dated to 17 ka.

In the MDS plot, we find that all of the newly and previously reported Epigravettian-associated individuals fall within the Villabruna cluster[4] (Fig. 1c). A series of f_4-symmetry statistics confirm that all the Epigravettian-associated individuals are cladal, and do not share excess affinity with any local (Paglicci 12) or non-local preceding ancestries (Goyet Q116-1, Kostenki 14, Mal'ta 1 or Věstonice) (Supplementary Data 2.F). Moreover, none of the Epigravettian-associated individuals have more affinity to southern European than to central-eastern European Gravettian-associated groups, as shown by f_4(Mbuti, Epigravettian-associated individual/group; Věstonice, Paglicci 12) that is consistent with 0 (Supplementary Data 2.G).

Next, we investigated the genetic relationships between Epigravettian-associated individuals across the Italian peninsula, by reconstructing a phylogeny based on a matrix of pairwise f_2 genetic distances (Fig. 3a and Supplementary Fig. 9) and testing the relative affinity among them using f_4-statistics in the form f_4(Mbuti, Epigravettian A; Epigravettian B, Epigravettian C) (Supplementary Data 2.E). The inferred topology reveals a phylogeographic pattern irrespective of individual ages. In particular, the 13 ka Pradis 1 individual from northeastern Italy represents the most basal lineage compared to all other Epigravettian-associated individuals, including the older Tagliente 2 and Villabruna genomes from northern Italy (group 1). Individuals from northwestern Italy (Arene Candide 16), central Italy (Continenza) and Sicily fall on a phylogenetically more derived branch (group 2), which further diversified into a branch composed of Sicilian hunter-gatherers only (group 3). Within Sicily, the 14 ka Oriente C individual shows higher affinity with the much younger but geographically

closer 10 ka Uzzo group[30] ($n = 2$) than with the almost contemporaneous San Teodoro 2 individual from eastern Sicily.

Fig. 3: The population substructure and diversity of Epigravettian-associated groups in southern Europe.


a
b
Latitude
Longitude
47.5
45.0
42.5
40.0
37.5
9° E
12° E
15° E
18° E
Pradis 1 (13 ka)
Villabruna (14 ka)
Tagliente 2 (17 ka)
Arene Candide 16 (13 ka)
Continenza (9–12 ka)
Paglicci 12 (29 ka)
Ostuni 1
(28 ka)
Oriente C (14 ka)
Uzzo
(10 ka)
San Teodoro 2 (13 ka)
Group 1
Group 2
Group 3
141
137
145
7
130
7
40
11
17
105
96
125
Pairwise mismatch rate
0.24
0.22
0.20
0.18
0.16
Gravettian
Group 1
Group 2
Group 3

a, Population structure among Epigravettian-associated populations revealed by a neighbour-joining tree based on pairwise f_2 genetic distances. Branch labels show unique drift lengths; black dots refer to individuals with newly generated data and white dots refer to previously published genomes; the position of each node does not imply the location where the split took place. **b**, Population diversity shown by PMR between individuals in different groups. The grouping of Epigravettian-associated populations is shown with the same colour in **a**. The black-outlined diamond marked out in the Gravettian group shows the PMR between the two Gravettian-associated individuals from southern Italy (Paglicci 12 and Ostuni 1). In the box plot the centre line is the median, box bounds delineate the interquartile range and whiskers extend to maximum and minimum values, excluding outliers. The sample size of individual pairs included in each group is reported in Supplementary Data 3.A.

Finally, we estimated the genetic diversity of Epigravettian-associated individuals in the dataset by calculating both pairwise mismatch rates (PMR) on pseudo-haploid genotypes and individual heterozygosity levels on pseudo-diploid genotypes (Supplementary Data 3.A). Compared with the genetic diversity observed among all analysed Gravettian-associated groups, Epigravettian-associated individuals show significantly lower amounts of genetic diversity (two-tailed *t*-test, $P < 0.001$) (Fig. 3b). Moreover, we reveal a north-to-south decrease in genetic diversity among the Epigravettian-associated groups, with the highest PMR and heterozygosity values found in northern Italian individuals (group 1), intermediate in western and central Italian individuals (group 2) and the lowest in Sicilian individuals (group 3) (Fig. 3b). A similar pattern is observed through the analysis of ROH segments (Extended Data Fig. 5 and Supplementary Information, section 9). We detect the highest amount of ROHs in Epigravettian-associated individuals from Sicily, who carry an extreme amount of more than 200 cM of short ROHs (4–8 cM). This suggests a very small recent effective population size, estimated to be in the order of around 70 individuals (Supplementary Table 2), causing the low genetic diversity in Sicilian Epigravettian hunter-gatherers.

To summarize, our results highlight a genetic turnover in the Italian peninsula of the Gravettian-associated Věstonice cluster by the Epigravettian-associated Villabruna cluster that might correlate with

discontinuities observed in the archaeological record[31]. We show that all analysed Epigravettian-associated individuals carry a homogeneous Villabruna ancestry, with the intra-group genetic structure mainly determined by their geographical, and not temporal, distribution. The phylogenetic reconstruction of Epigravettian-associated genomes, with Pradis 1 diverging more deeply than all others, indicates that the turnover took place much earlier than 17 ka—the date of the more derived Tagliente 2 genome. This, together with the evidence of Villabruna ancestry in El Mirón 19 ka, further suggests that this genetic discontinuity could be the result of palaeogeographic and palaeoecological transformations connected to the LGM[32], rather than to the Bølling–Allerød warming period[4,15] (14.7–12.9 ka). In addition, our phylogeographic analysis points to northeastern Italy as the possible entry point of the Epigravettian-associated gene pool in the Italian peninsula. This finding, in conjunction with the genetic affinity of the Villabruna cluster to ancient and present-day Near Eastern ancestries[4,15,16] (Supplementary Information, section 8, Supplementary Fig. 18 and Supplementary Data 2.O), suggests the Balkans as a source of the incoming Epigravettian-associated population. The LGM could thus have created a corridor south of the Alps for east-to-west human movements that genetically connected hunter-gatherer populations from the Balkans to Iberia, possibly also via dispersals along existing lower-sea-level coasts[32].

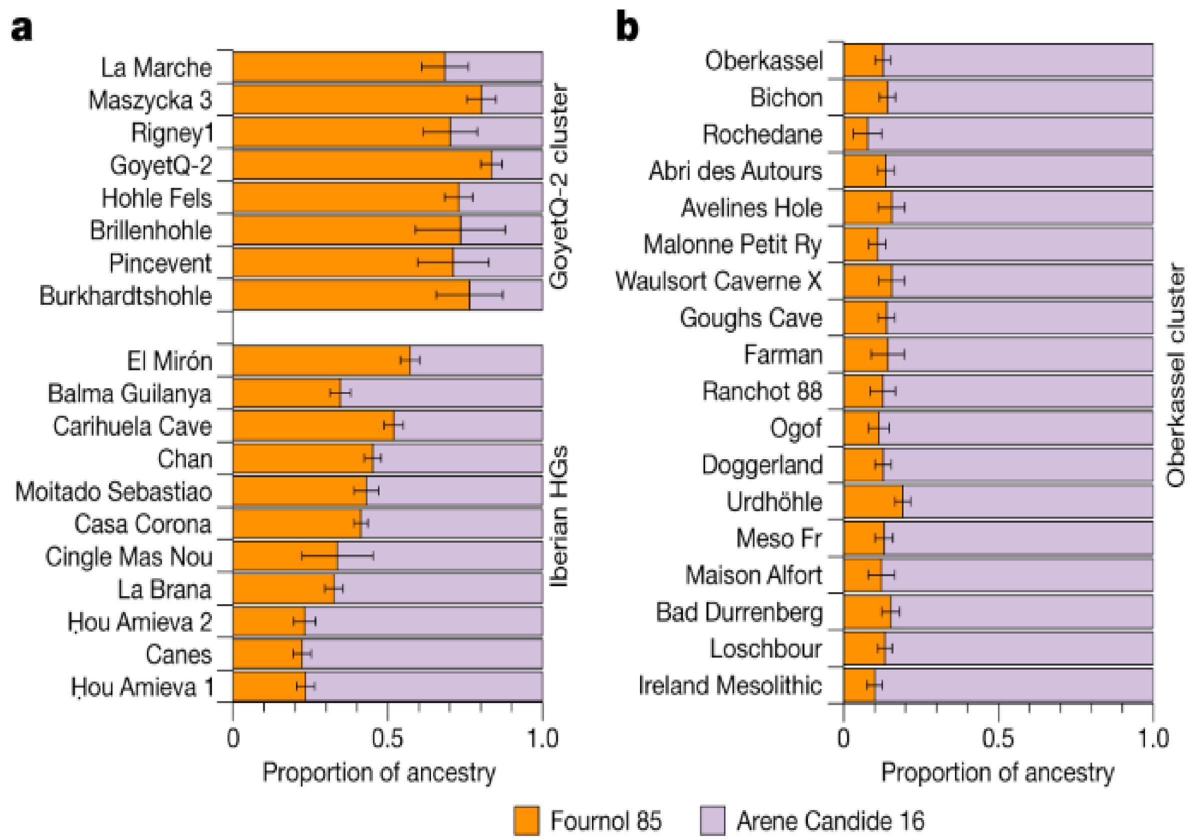
Post-LGM in western and central Europe

The Magdalenian culture was widely distributed in southwestern, western and central Europe after the LGM[33]. Despite this wide geographical range, it is not clear whether different groups associated with this culture originated from a common source population and how those groups were genetically related to each other. Previous studies identified two different genetic compositions in Magdalenian-associated individuals—the GoyetQ2 cluster including central-western European genomes dated to around 15 ka (from France, Belgium and Germany), and the ancestry of the El Mirón individual from Spain[4,14] from around 19 ka. Both of these ancestries carry a genetic component distantly related to the Goyet Q116-1 individual dated to 35 ka, with the Iberian individual also showing an affinity to the Villabruna cluster[4,14]. By co-analysing previously published data with our newly

reported genomes associated with the Magdalenian from La Marche (18 ka) and Pincevent (15 ka) in western and northern France, respectively, and Maszycka (18–16 ka) in southern Poland, we confirm that the Goyet Q116-1 ancestry survived in all studied Magdalenian-associated genomes besides in Gravettian and Solutrean-associated individuals from southwestern and western Europe (Fig. 1). Notably, the Fournol ancestry provides a better proxy than Goyet Q116-1 for the genetic component found in the GoyetQ2 cluster and in El Mirón (Supplementary Data 2.H). However, using f_4-statistics, we show that all Magdalenian-associated individuals, and not only El Mirón, carry Villabruna-related ancestry when compared to the Fournol cluster (Supplementary Data 2.H). This affinity is even stronger towards Epigravettian-associated individuals from western and central Italy and Sicily (group 2 and group 3, respectively) than to those from northern Italy (group 1) (Supplementary Data 2.F).

We thus modelled individuals belonging to the GoyetQ2 cluster and El Mirón as a mixture between the Fournol 85 and Arene Candide 16 genomes as proxies to represent the Fournol and Villabruna ancestries, respectively, in Magdalenian-associated groups (Fig. 4a). Besides El Mirón, who has around 43% Villabruna ancestry, all other Magdalenian-associated individuals have a lower proportion of this component (19–29%) and can thus be assigned to the GoyetQ2 cluster (Fig. 4a and Supplementary Data 3.C). This is further validated by f_4-statistics of the form f_4(Mbuti, Arene Candide 16; Goyet Q-2, Magdalenian-associated individuals), which is significantly positive only for El Mirón, whereas all other tested individuals and Goyet Q-2 are symmetrically related with respect to Arene Candide 16 (Supplementary Fig. 26 and Supplementary Data 2.H).

Fig. 4: Ancestry modelling of post-19 ka individuals in southwestern, western and central Europe.



a,**b**, The ancestries of individuals in the GoyetQ-2 cluster and Iberian hunter-gatherers (HGs) (**a**) and individuals in the Oberkassel cluster (**b**) were modelled using qpAdm, with Fournol 85 and Arene Candide 16 representing the Fournol and Villabruna ancestries, respectively. The length of the colour bar shows the proportion of each ancestry. The error bar shows the s.e.m. of estimates from 5-cM-block jackknife analysis. Details of the modelling are provided in Supplementary Data 3.C.

Our analyses demonstrate that the Fournol cluster is a better source for Magdalenian-associated genomes than Goyet Q116-1. Therefore, most of the ancestry found in these post-LGM individuals probably traced back to Gravettian-associated groups from western and southwestern Europe. The genetic affinity to the Villabruna ancestry is present in El Mirón and in Magdalenian-associated individuals from western and central Europe. This suggests that genetic links between southern and southwestern European hunter-gatherers around the time of the LGM extended north of the Pyrenees. The resulting GoyetQ2 cluster includes individuals spanning from western France to Poland in the period between 18 and 15 ka. Therefore,

contrary to previous suggestions[34], this demonstrates that the post-LGM diffusion of the Magdalenian was indeed associated with northward and northeastward population expansions from western Europe[35].

Post-14 ka to Neolithic

Previous studies have shown that two main hunter-gatherer ancestries were predominant across most parts of Europe after around 14 ka—that is, the western hunter-gatherer (WHG) ancestry, related to the Villabruna cluster, and the eastern hunter-gatherer (EHG) ancestry, showing affinity to both the Villabruna and the ancient north Eurasian (ANE) ancestry found in Upper Palaeolithic Siberian individuals[4,36]. Hunter-gatherers carrying an admixed WHG/EHG genetic profile have been sequenced from various regions of northern and eastern Europe, raising the question of how these two types of ancestries formed and interacted with each other through time and space[37,38,39,40].

In the MDS plot (Fig. 1c) and a west Eurasian principal component analysis (PCA) (Extended Data Fig. 6 and Supplementary Fig. 14), most post-14 ka individuals from western and central Europe fall close to the WHG cluster and those from eastern Europe close to the EHG cluster, whereas the Tutkaul 1 individual from central Asia falls close to the ANE-related group. The two 14 ka Oberkassel individuals mark the earliest presence of WHG ancestry north of the Alps, which we therefore rename the Oberkassel cluster (hereafter, Oberkassel cluster or ancestry), using the name of the oldest reported individual to date carrying such ancestry with more than one-fold coverage, for consistency[4]. On the basis of f_4-statistics, we find that individuals assigned to the Oberkassel cluster are closer to the Arene Candide 16 genome than any other Epigravettian-associated group from Italy (Supplementary Data 2.F). Moreover, the Oberkassel cluster carries both Villabruna ancestry and a contribution from GoyetQ2 ancestry (Supplementary Data 2.J). This was confirmed with qpAdm, in which we could model all individuals from the Oberkassel cluster as a broadly constant mixture of approximately 75% Arene Candide 16 and 25% Goyet Q-2 (or 90% Arene Candide 16 and 10% Fournol 85) (Fig. 4b and Supplementary Data 3.C). The observation that post-14 ka individuals from

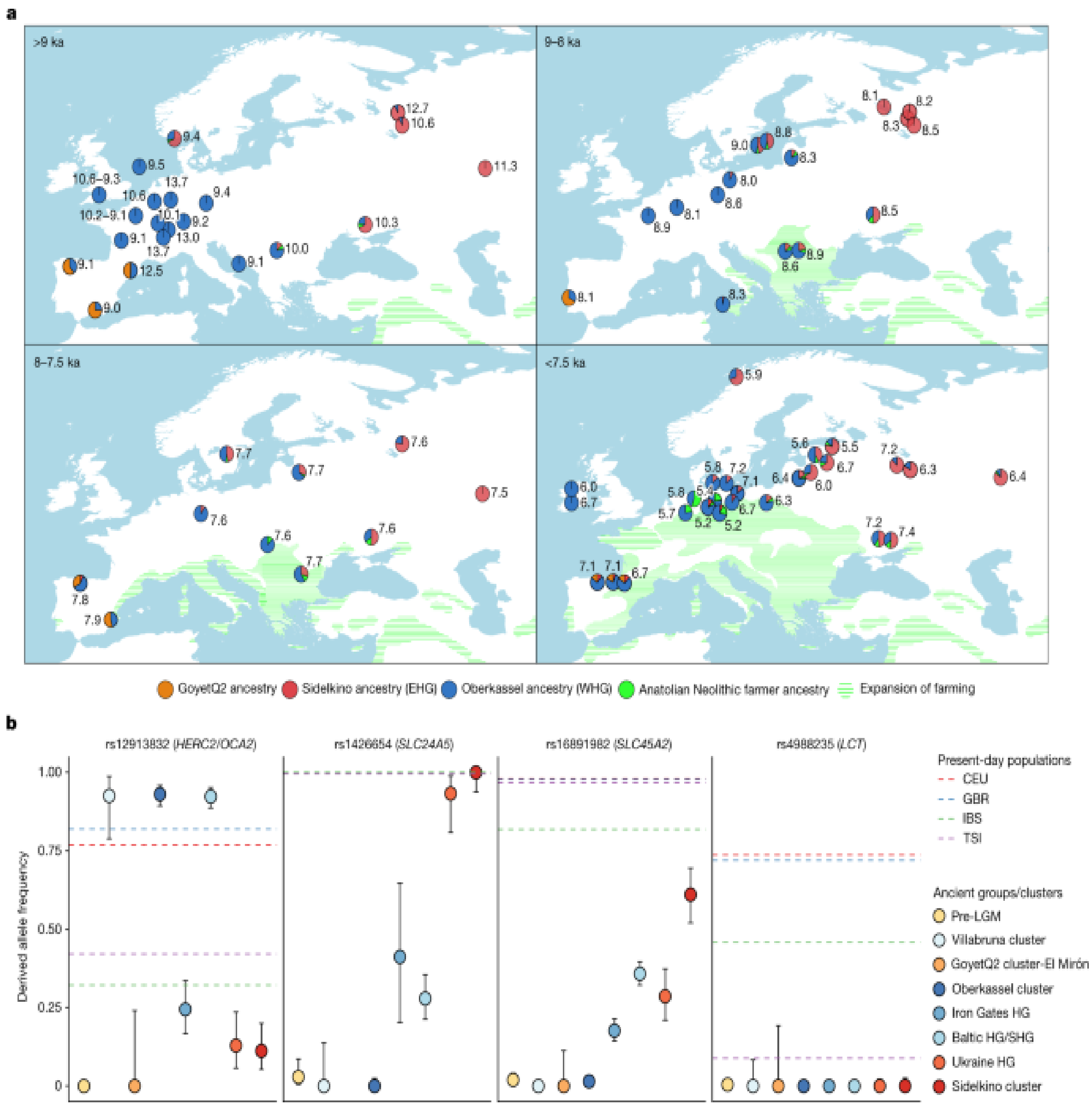
western and central Europe and also from Britain[41] carry a homogeneous genetic makeup instead of displaying repeated local admixtures with GoyetQ2 ancestry implies that the Oberkassel-ancestry profile was already largely formed before its dispersal. This is in sharp contrast to the genetic history of Iberian hunter-gatherers, where the spread of the Villabruna/Oberkassel ancestry involved multiple local admixture events with groups carrying high proportions of GoyetQ2 ancestry[14] (Fig. 4 and Supplementary Data 3.C). The long-lasting genetic continuity in Iberia is also reflected in the preservation until the Mesolithic of Y-chromosome haplogroup C, which was predominant in pre-LGM groups but rarely found after the LGM in other parts of Europe (Extended Data Figs. 1 and 2).

Using f_4-statistics and qpAdm, we confirm that EHG populations in eastern Europe are a mixture of Villabruna/Oberkassel and ANE ancestries (Supplementary Information, section 11 and Supplementary Data 2.K). F_4-statistics also show that the approximately 8.2 ka Yuzhniy Oleniy Ostrov group from Karelia in western Russia formed by 19 genomes has comparable or lower affinity to Villabruna ancestry than all the other EHG groups (Supplementary Data 2.K). The oldest individual revealing an indistinguishable genetic profile from the Yuzhniy Oleniy Ostrov group is the 11 ka Sidelkino individual from Samara in western Russia[42]. For consistency with the previously discussed nomenclature, we rename the EHG ancestry as the Sidelkino cluster (hereafter, Sidelkino cluster or ancestry). The genetic distinction between the Oberkassel and Sidelkino clusters is also clearly noticeable in the diversity of uniparentally inherited markers, as the Oberkassel cluster is dominated by mtDNA haplogroup U5 and Y-chromosome haplogroup I, whereas individuals from the Sidelkino cluster show a higher frequency of mtDNA haplogroups U2, U4 and R1b, and carry uniquely Y-chromosome haplogroups Q, R and J (Extended Data Figs. 1 and 2).

We then attempted to model 250 published and newly reported hunter-gatherers dated to 14–5 ka using qpAdm as a mixture of Oberkassel, Sidelkino, GoyetQ2 ancestries, and an ancestry maximized in Anatolian Neolithic farmers (ANF), as a considerable portion of the sequenced hunter-gatherer genomes date after around 8 ka, when ANF ancestry started spreading across Europe. Our results show that the contact zone and the

admixture patterns between the Oberkassel and Sidelkino ancestries changed over time (Fig. 5). Between 14 and 8 ka, all hunter-gatherers in western and central Europe carried only Oberkassel ancestry, with no detectable contribution from the Sidelkino cluster. Further north and east, individuals from the Baltics (Baltic HG), Scandinavia (SHG), the Balkans (Iron Gates HG) and Ukraine (Ukraine HG) already carried an Oberkassel/Sidelkino admixed ancestry[38,40] before 8 ka. In addition, those groups also carry affinity to ANF suggesting more complex genetic processes behind their demographic history[16]. Moreover, two of the oldest published groups from western Russia belonging to the Sidelkino cluster—Peschanitsa (13 ka)[43] and the newly reported Minino individuals (11 ka)—showed extra affinity to the Oberkassel cluster, possibly owing to variability in this ancestry proportion during the initial formation phase of the Sidelkino-ancestry profile. Using DATES software, we estimated the admixture between Villabruna/Oberkassel and ANE ancestries in these old Sidelkino-cluster-related individuals to around 15–13 ka (Extended Data Fig. 7 and Supplementary Table 3), which coincides roughly with the first appearance of the Oberkassel ancestry in central Europe. This raises the possibility that the replacement by the Oberkassel cluster and the formation of the Sidelkino cluster might have been the result of population expansions influenced by the abrupt warming during the Bølling–Allerød interstadial[4,24].

Fig. 5: Ancestry modelling of hunter-gatherers from 14–5.2 ka and their allele frequencies on phenotypic SNPs.



a, The genetic ancestry of hunter-gatherers dated between 14 ka and 5.2 ka modelled using qpAdm, with Oberkassel, Yuzhniy Oleniy Ostrov, Goyet Q-2 and Neolithic farmers from present-day Turkey (Barcın, Menteşe and Boncuklu sites) representing Oberkassel (WHG) (blue), Sidelkino (EHG) (red), GoyetQ2 (orange) and Anatolian Neolithic farmer (green) ancestries, respectively. The average calibrated date is shown, with pie charts indicating the estimated proportion of ancestry for each group or individual. Details of the modelling are provided in Supplementary Data 3.E,F. The expansion of farming by 9, 8, 7.5 and 7 ka is shown as green shades. Adapted from https://doi.org/10.5281/zenodo.5903165 (CC BY 4.0). **b**, Allele frequencies of different hunter-gatherer groups (coloured dots) on four SNPs related to

skin colour (*SLC24A5* and *SLC45A2*), eye colour (*HERC2*/*OCA2*) and lactase persistence (*LCT*). Dots are maximum likelihood estimates and error bars show 95% confidence intervals of the derived allele frequencies (*n*, the number of individuals in each group, is provided in Supplementary Data 3.G). Dashed lines show the frequencies estimated for the indicated present-day 1000 Genomes Project populations (CEU, Utah residents of northern and western European ancestry; GBR, British; IBS, Spanish; TSI, Tuscan)[37]. Details on the allele frequency estimates are provided in Supplementary Information, section 12, Supplementary Figs. 27 and 28 and Supplementary Data 3.G.

From around 8 ka, we begin to observe admixture events with Sidelkino ancestry in central Europe. This is first detected in an individual from Gross Fredenwalde in northeastern Germany and reaches around 10% in most European hunter-gatherer individuals thereafter (Extended Data Fig. 8). Soon after 8 ka, Sidelkino ancestry was absent in eastern Spain but it had already reached northern Iberia alongside an increase in Oberkassel ancestry (Fig. 5). Conversely, additional Oberkassel ancestry is identified in eastern Europe by at least 7.5 ka in newly generated genomes from Minino I and Yazykovo from the upper Volga region, whereas a 1,000-years-older individual from Minino I did not have this genetic component. Considering a freshwater reservoir signal in the upper Volga region making radiocarbon dates on human remains appear up to about 500 years older than their true age[44], there could be an interval of more than 1,000 years between the first evidence of admixture in central European hunter-gatherers with Sidelkino ancestry and eastern European hunter-gatherers with Oberkassel ancestry. However, additional genomes intermediate in time and space are needed to assess whether those two admixture events were independent or part of a common demographic process.

After 7.5 ka, as ANF ancestry had reached regions north of the Alps, individuals carrying a hunter-gatherer genetic profile were primarily restricted to the northern fringes of Europe (Fig. 5). In this period, the Oberkassel-ancestry admixture spread further east, reaching Samara by around 6.5 ka, and an increase in Sidelkino ancestry was detected in hunter-gatherers from the Baltic region, which was previously associated with the transition from the Narva culture to the Comb Ceramic culture[38,39]

(Extended Data Fig. 8). In central Europe, admixture with ANF ancestry became highly common but not ubiquitous, indicating the co-existence of hunter-gatherer and farmer societies without admixing for several hundred years. The youngest individual carrying large portions of hunter-gatherer ancestry in the analysed dataset is from Ostorf in northern Germany, dated to around 5.2 ka (>90% Oberkassel cluster plus Sidelkino-cluster components) (Supplementary Data 3.F). Individuals at this site might mark one of the last occurrences of such high levels of hunter-gatherer-related ancestries, just centuries before the emerging European Bronze Age.

On the basis of PCA and outgroup f_3-statistics, the Neolithic Tutkaul 1 individual from Tajikistan is closely related to Upper Palaeolithic individuals from south-central Siberia (Afontova Gora 3 (AG3) and Mal'ta 1), and roughly contemporaneous West Siberian hunter-gatherers (Tyumen and Sosnoviy), both carrying high proportions of ANE ancestry[45] (Fig. 1c and Extended Data Fig. 6). We tested the affinity of Tutkaul 1 to worldwide ancient and modern populations relative to AG3. Contrary to West Siberian hunter-gatherers, Tutkaul 1 does not carry an extra eastern Eurasian ancestry, but shows affinity to Iranian Neolithic farmers and some younger populations from Iran and the Turan region (Supplementary Data 2.L). Conversely, individuals in the Sidelkino cluster are genetically closer to AG3 than Tutkaul 1. This suggests that the newly reported Neolithic individual from central Asia carries an ancestry that might be a good proxy for the ANE-related contribution to Iran and the Turan region[45] from around 5.5 ka but not to roughly contemporaneous hunter-gatherers from eastern Europe.

In sum, we describe the formation and interaction between the Oberkassel and Sidelkino clusters, the two main hunter-gatherer ancestries present in Europe from 14 ka onwards. The genomic similarity of the Oberkassel cluster to Arene Candide 16 in northwestern Italy might imply that Epigravettian-associated ancestry spread from the south to central Europe passing through the western side of the Alpine region. The Sidelkino ancestry also emerged around 14 ka with its first direct evidence in eastern Europe[43] dated to 13 ka. The increasing level of admixture between distinct hunter-gatherer populations from around 8 ka onwards indicates an intensified mobility of those forager groups. This might have been in part

triggered by the concomitant expansion of Neolithic farmers across Europe and/or by environmental factors, such as the climatic event around 8.2 ka, the largest abrupt cooling in the northern hemisphere during the Holocene epoch[46,47].

Phenotypically relevant variants

Leveraging the substantially increased sample size, we investigated genetically distinct hunter-gatherer groups for allele frequencies at selected loci that are known to be associated with specific phenotypic traits in present-day Europeans (Fig. 5b and Supplementary Figs. 27 and 28). Consistent with previous findings, none of the analysed groups show the derived allele at SNP rs4988235 on the *LCT* gene, which is responsible for lactase persistence. As previously hinted[37], we find a large frequency variation in alleles related to skin and eye pigmentation among post-LGM hunter-gatherer groups. For the SNP associated with light eye colour (*HERC2/OCA2* (rs12913832)), individuals from the Villabruna cluster, Oberkassel cluster, Baltic HG and SHG groups show high frequencies of the derived allele (>90%), which is responsible for the green or blue eye phenotype, whereas Sidelkino cluster, Ukraine HG and Iron Gates HG groups show low occurrence of this allele (10–25%). Instead, for the two SNPs associated with skin colour (*SLC24A5* (rs1426654) and *SLC45A2* (rs16891982)), Sidelkino cluster and Ukraine HG groups show a higher frequency (>90% for *SLC24A5* and 29–61% for *SLC45A2*) of the derived alleles related to light skin colour, compared with Oberkassel and Villabruna clusters, where those alleles are almost completely absent (<1%). On the basis of the genetic variation of present-day Europeans, this could imply phenotypic differences between post-14 ka hunter-gatherer populations across Europe, with individuals in the Oberkassel cluster possibly exhibiting darker skin and lighter eyes, and individuals in the Sidelkino cluster possibly lighter skin and darker eye colour.

Discussion and conclusions

The data generated in this study enabled us to investigate genomic transformations of and interactions between Eurasian hunter-gatherers at

high resolution (Extended Data Fig. 9). We provide five novel insights into the genomic history of hunter-gatherer populations over a time span of 30,000 years from the Upper Palaeolithic to the Neolithic.

First, we show that individuals associated with the Gravettian culture across Europe were not a biologically homogeneous population. Culturally, however, we see both widespread general tendencies, such as weaponry and some portable art[48], and other aspects that have a more regional character, such as mortuary practices (Supplementary Information, section 13), various originalities in lithic and hard organic materials tool kits and adornments[20,21]. The ancestry found in individuals associated with the preceding Aurignacian culture from central Europe (GoyetQ116-1 ancestry) gave rise to Gravettian-associated individuals from western and southwestern Europe. This derived ancestry—the Fournol cluster—survived during the LGM in Solutrean-associated individuals, possibly within the Franco-Cantabrian climatic refugium[25], leading to later populations associated with the Magdalenian culture (GoyetQ2 cluster and El Mirón). Conversely, the ancestry found in pre-30 ka eastern European individuals (Kostenki cluster and Sunghir group) contributed to Gravettian-associated individuals from central and southern Europe (Věstonice cluster), the latter without descendants retrieved in post-LGM populations from those regions.

Second, the ancestry of individuals associated with the Epigravettian culture (Villabruna cluster), which was found to genetically connect European and Near Eastern hunter-gatherers, reached southern Europe well before the transition between the Early and Late Epigravettian[4,15] and possibly as early as the Gravettian–Epigravettian transition. A phylogeographic reconstruction of different lineages carrying this ancestry further suggests its entry point into northeastern Italy from the Balkans followed by a north-to-south expansion into the Italian peninsula alongside a population decline through sequential bottlenecks.

Third, Magdalenian-associated individuals not only from Iberia but also from the rest of Europe carry Epigravettian-associated ancestry (Villabruna cluster). Genetic analyses of western European individuals associated with the preceding Badegoulian culture might provide clues on the processes that

led to the formation of the GoyetQ2 cluster. As inferred from the archaeological record[35], the spread of the Magdalenian across Europe is linked to southwestern to northern and northeastern post-LGM population expansions and not to movements from southeastern refugia[34].

Fourth, we extend the finding of a large-scale genetic turnover as early as 14 ka in central and western European hunter-gatherers associated with multiple techno-complexes—Federmesser, Azilian and other Final Palaeolithic groups[4]—despite considerable technological continuity with the preceding late Magdalenian. This broadly distributed ancestry (the Oberkassel cluster (also known as WHG)) is most closely related to an Epigravettian-associated individual from northwestern Italy, suggesting that its expansion into continental Europe might have started from the west—and not the east—side of the Alps. Moreover, the almost complete genetic replacement of the Magdalenian-associated gene pool raises the hypothesis that parts of Europe were differentially populated during the abrupt climatic variation starting around 14.7 ka with the Bølling–Allerød warming period, creating areas where southern European populations could expand. This might also explain the genetic uniformity of the Oberkassel cluster across large parts of western Eurasia but genomic data from between 15 and 14 ka is needed to understand the exact dynamics of this turnover.

Fifth, the Oberkassel ancestry in western and central Europe and the Sidelkino ancestry in eastern Europe remained largely isolated for almost 6,000 years until genetic interactions were first observed—around 8 ka in northeastern Germany, possibly associated with cultural exchanges along the Baltics[49] and around 7.5 ka in the upper Volga region, possibly linked to the spread of pottery in the region[50].

In conclusion, our study reveals that western and southwestern Europe served as climatic refugia for the persistence of human groups during the coldest phase of the last Ice Age whereas populations in the Italian peninsula and the eastern European plain were genetically overturned, challenging the role of these regions as glacial refugia for humans. The incoming Villabruna ancestry later became the most widespread hunter-gatherer ancestry across Europe. Further palaeogenomic studies on Upper Palaeolithic individuals from the Balkans will be essential for understanding whether southeastern

Europe represents the source of the Villabruna ancestry and a climatic refugium for human populations during the LGM.

Note added in proof: A companion paper[51] describes genome-wide data of a 23,000-year-old Solutrean-associated individual from southern Iberia that extend the evidence of genetic continuity across the LGM in southwestern Europe.

Methods

Archaeological sampling

The ancient human specimens analysed in this work derive from multiple scientific collaborations. All remains were sampled with the approval of the institutions responsible for the analysis of archaeological material. This was achieved through collaboration with local curators and scientists from the countries where the skeletal material is preserved and who are listed among the authors of this study. The responsible co-authors for the material from each archaeological site are listed in Supplementary Information, section 1.

The analysed individuals span from the Upper Palaeolithic to the Neolithic. While terms such as lithic industry, techno-complex, prehistoric tradition, and so on might be more appropriate to refer to the various associated chrono-cultural subdivisions, they concern different levels of discussion and are not applicable to all contexts investigated here. Therefore, the broader terms ‘archaeological culture’ or simply ‘culture’ are used here to refer to archaeologically defined material cultures without implying links to modern anthropological and/or ethnographical concepts of culture.

Radiocarbon dating

We report 47 new radiocarbon dates performed on skeletal elements of 40 individuals by the Curt-Engelhorn-Zentrum Archaeometrie in Mannheim (MAMS, $n = 29$), Center for Isotope Research, University of Groningen (GrA and GrM, $n = 5$), University of Aarhus (AAR, $n = 3$), Beta Analytics (Beta, $n = 2$), Zürich (ETH, $n = 3$), International Chemical Analysis (ICA, $n = 2$), Natural History Museum in Paris (Echo Lab, $n = 1$) and Vilnius

(FTMC, $n = 2$) (Supplementary Data 1.A). The dates were calibrated using OxCal 4.4[52] with calibration curve IntCal20 at 95.4% probability[53] and when multiple dates were available for the same individual we used the function R_Combine to combine them[52]. We did not correct the calibrated dates for marine or freshwater reservoir effects but, when available, we report individual stable isotope values ($\delta^{15}N/\delta^{13}C$ and C:N ratio) in Supplementary Data 1.A to evaluate the potential impact of such reservoir effects.

Ancient DNA processing

The human remains were processed in dedicated laboratories at the Max Planck Institute for the Science of Human History in Jena (Germany), University of Tübingen (Germany), University of Florence (Italy), Leiden University Medical Center (the Netherlands) and University of Tartu (Estonia). Human bones and teeth were sampled in clean room facilities to minimize the inclusion of modern human DNA contamination during this procedure. DNA was extracted from the generated bone or tooth powder following established protocols. A subset of samples (GER002 and GER003) were pre-treated with a washing step to reduce surface contamination[54]. A negative and cave bear positive controls were included. For the DNA lysis, a solution of 900 µl EDTA, 75 µl H_2O and 25 µl proteinase K was added. In a rotator, samples were digested for at least 16 h at 37 °C, and for pre-treated samples this was followed[55] by an additional hour at 56 °C. The suspension was then centrifuged and transferred into a binding buffer as previously described[56]. To bind DNA, silica columns for high volumes (High Pure Viral Nucleic Acid Large Volume Kit (Roche)) were used. After 2 washing steps using the manufacturer's wash buffer, DNA was eluted in TET (10 mM Tris, 1 mM EDTA and 0.05% Tween) in two steps for a final volume of 100 µl. After DNA lysis, a subset of samples was extracted using silica-coated magnetic particles on an automated liquid handling system (Agilent Technologies Bravo NGS Workstation)[57]. Double-stranded DNA libraries were built from 25 µl of DNA extract, without the presence of uracil DNA glycosylase (ds_nonUDG) or in the presence of uracil DNA glycosylase (ds_halfUDG), following a double-stranded 'UDG-half' library preparation to reduce, but not eliminate, the amount of deamination-induced damage towards the ends of ancient DNA (aDNA)

fragments[58]. Negative and positive controls were carried alongside each experiment. Libraries were quantified using the IS7 and IS8 primers[59] in a quantification assay using a DyNAmo SYBR Green qPCR Kit (Thermo Fisher Scientific) on the LightCycler 480 (Roche). Each aDNA library was double indexed[60] in 1–4 parallel 100 µl reactions using PfuTurbo DNA Polymerase (Agilent). The indexed products for each library were pooled, purified over MinElute columns (Qiagen), eluted in 50 µl TET and again quantified using the IS5 and IS6 primers[59] using the quantification method described above. The purified products were amplified in multiple 100 µl reactions using Herculase II Fusion DNA Polymerase (Agilent) following the manufacturer's specifications with 0.3 µM of the IS5/IS6 primers. After another MinElute purification, the product was quantified using the Agilent 2100 Bioanalyzer DNA 1000 chip. An equimolar pool of all libraries was then prepared for shotgun sequencing on Illumina Hiseq4000 platform using 75bp single-end reads for screening. Single-stranded DNA libraries were built from 30 µl of DNA extract in the absence of uracil DNA glycosylase (ss_nonUDG) followed by double indexing, using an automated version of the protocols described in[61] on the liquid handling system mentioned before. The single-stranded library of Cuiry Les Chaudardes 1 was produced with partial UDG treatment (ss_halfUDG)[62] (Supplementary Data 1.B).

DNA enrichment and sequencing

Both double-stranded and single-stranded libraries were further amplified with IS5/IS6 primers to reach a concentration of 200–400 ng/µl as measured on a NanoDrop spectrophotometer (Thermo Fisher Scientific). The libraries underwent shallow shotgun sequencing on an Illumina HiSeq 4000 instrument with 75 single-end-run cycles using the manufacturer's protocol, to evaluate the human endogenous DNA content and quality. Samples with a percentage of human DNA in shotgun data around 0.1% or greater were enriched for a set of 1,237,207 targeted SNPs (1240k capture) across the human genome[6]. mtDNA capture[63] was also performed for those libraries where mtDNA coverage was not high enough to assess mtDNA haplogroup and contamination. Illumina sequencing platforms were also used to sequence the 1240k and mtDNA captured libraries (Supplementary Data 1.B).

The de-multiplexed capture sequencing reads were cleaned and mapped to human reference genome hs37d5 using EAGER pipeline 1.92.55[64]. Within the pipeline, the adapters were removed by AdapterRemoval 2.2.0[65], reads were mapped with BWA 0.7.12 aln/samse algorithm[66], duplications were removed by DeDup 0.12.1 (https://github.com/apeltzer/DeDup) and damage patterns of each library were checked with mapDamage 2.0.6 and 2.0.9[67]. The deduplicated bam files were filtered using PMDtools 0.60[68] with a threshold of 3, to reduce potential modern DNA contamination based on postmortem DNA deamination. For ds_halfUDG libraries, we masked 2 bp from both ends of the reads with trimBam in bamUtil 1.0.13 (https://github.com/statgen/bamUtil) to remove the damaged sites.

The mitochondrial capture sequencing reads were cleaned by AdapterRemoval 2.2.0 to remove the adapters and reads with lengths below 30 bp. Then the cleaned reads together with cleaned reads from 1240k capture sequencing were mapped to human reference mitochondrial sequence NC_012920.1 with BWA 0.7.12 aln/samse algorithm (parameters –n 0.01, –l 16500) and realigned with CircularMapper[64]. The mapped reads from the same individual and library set-up were merged and duplications were removed with DeDup. Reads with a mapping quality below 30 were then filtered with samtools, and the consensus sequences were generated by Schmutzi[69].

Ancient DNA authentication and genotyping

The sex of each individual was determined by the ratio of sequencing coverages on sex chromosomes versus autosomes (Supplementary Data 1.C). Individuals with libraries showing signs of contamination were further tested using PMD-filtered bam files. Individuals with at least one library showing Y/Auto ratio > 0.2 were determined as male individuals, and with Y/Auto < 0.2 were determined as female individuals[4] (Supplementary Fig. 12).

The nuclear DNA contamination was estimated with several methods. We applied ANGSD 0.934[70] and hapCon[71] for libraries from male individuals, and applied contamLD[72] and a newly developed method that analyses contamination in ROH for female and male libraries (see Supplementary

Information, section 2 for a detailed description). The mtDNA contamination was estimated by Schmutzi (--notusepredC --uselength)[69] for all the libraries. Libraries showing a mitochondrial or nuclear contamination rate over 10% were considered substantially contaminated whereas those between 5 and 10% were considered marginally contaminated and were treated differently (details are provided in Supplementary Information, section 3).

The cleaned reads with base quality and mapping quality over 30 were piled up with mpileup in SAMtools 1.3[73] on the 1240k targeted sites. For contaminated libraries we used the PMD-filtered bam files as the input for genotyping. Then pseudo-haploid genotypes were called using pileupCaller 1.4.0.2 (https://github.com/stschiff/sequenceTools) under random haploid calling mode. For ds_halfUDG libraries, we called genotypes on all targeted sites from 2bp-masked bam files; for ds_nonUDG libraries, we called genotypes on transversion sites only; for ss_nonUDG libraries, we called genotypes with single-strand mode, which ignores forward reads at C/T polymorphisms and reverse reads at G/A polymorphisms.

Then we merged the genotypes from different libraries of the same individual, by randomly picking alleles from available genotype calls, using a custom script. After merging, individuals with less than 6,000 SNPs on 1240k sites were excluded from further analysis because of low coverage. We also genotyped a selection of previously published individuals with the same approach (Supplementary Data 1.G)[2,4,30,74,75,76,77,78]. Then we combined our newly generated genotypes with published genotypes from ancient and modern individuals from AADR v42.4 (Allen Ancient DNA Resource (https://reich.hms.harvard.edu/allen-ancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dna-data) version 42.4) for downstream analysis[1,4,7,14,16,23,36,37,38,39,40,42,43,45,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94].

For individual heterozygosity calculation, we also called pseudo-diploid genotypes from each library, using pileupCaller 1.4.0.3 under random diploid calling mode and the same strategy for different types of libraries as pseudo-haploid genotype calling.

Uniparental markers

The mitochondrial haplogroups were determined using HaploGrep 2[95], based on the consensus sequences generated from Schmutzi inspected for each sample at increasing quality filters (from q0 to q20). Inconsistent haplogroup assignments were manually verified as indicated[24] (Supplementary Data 1.L). For phylogenetic reconstruction (Extended Data Fig. 1) we used MUSCLE (-maxiters 2)[96] to create a multiple genome alignment of previously published sequences and newly reported mtDNA consensus sequences with q20 according to defined thresholds (minimum average coverage >5-fold, contamination estimate <20%, HaploGrep 2 haplogroup assignment consistent with manual assignment). We built a Maximum Parsimony tree with 103 mtDNA sequences plus an African sequence as the outgroup (not shown) after the removal of individuals younger than 6.5 ka and mtDNAs from the same site with an identical placement. The tree was calculated on 16,528 positions (partial deletion 95%) and with 500 bootstrap iterations using MEGA10[97].

To determine the Y-chromosome haplogroups of male individuals, we genotyped the Y-chromosome reads using a Y-SNP list (v.15.73) from the International Society of Genetic Genealogy (ISOGG) dataset, ignoring C-to-T and G-to-A transitions on the forward and reverse reads, respectively. This procedure allowed us to manually traverse the ISOGG Y-Haplogroup Tree, checking in a semi-automatic way which positions were covered. This process allowed us to assign an ancestral or derived haplogroup for covered branches, and to make corrections to calls in cases where, for instance, a more derived haplogroup was called because of residual ancient damage (C-to-T or G-to-A mismatches) in terminal read positions at diagnostic SNPs[98] (Supplementary Data 1.A). For the placement of individuals onto a Y-chromosomal phylogenetic tree (Extended Data Fig. 1), we used pathPhynder[99] based on the tree from Karmin et al.[100]. We used the default posterior threshold of 0.01, and mapping and sequencing quality cutoffs of 30. We then removed samples with less than 0.04X coverage (calculated on the mappable, non-recombining region of the Y chromosome[98]) to avoid arbitrarily placing low-coverage samples at the root of major haplogroups. This results in a tree with 57 newly reported and previously published ancient individuals while present-day sequences are collapsed in the major

Y-chromosome haplogroups (the most basal lineages are not shown). The tentative placements of low-coverage ancient individuals based on their haplogroup assignment (Supplementary Data 1.A) are indicated with arrows on the respective branches.

Biological relatedness and population diversity

The analysis of biological relatedness was performed by calculating relatedness coefficient (*r*) based on PMR on the autosomal SNPs (Supplementary Data 1.F and Supplementary Information, section 4). The baseline of each population was determined using the average heterozygosity rate of individuals estimated from pseudo-diploid genotypes (Supplementary Data 1.E).

The ROH segments in hunter-gatherer genomes were identified using hapROH[101]. As recommended, we analysed individuals with over 400,000 SNPs called on the 1240k panel[101] and we called ROH longer than 4 cM (Supplementary Data 3.B). The effective population sizes (N_e) were then estimated using a maximum likelihood method, after filtering individuals with a signal of close-kin inbreeding (individuals with at least 50 cM of their genome in ROH spanning >20 cM) (Supplementary Information, section 9).

Population genetics analysis

The PCA was carried out by smartpca in EIGENSOFT 6.0.1[102], with modern individuals used for calculation and all the ancient individuals projected on the calculated PCs. The "lsqproject: YES" parameter was used to minimize the effect of missing data in ancient individuals. The PCA was calculated with 1379 individuals from 87 western Eurasian modern populations on the 1240k_HO dataset, which was intersected between the 1240k and Human Origins datasets (Supplementary Data 1.K).

The MDS analysis showing the genetic affinity among European hunter-gatherers was based on the distance matrix derived from outgroup f_3-statistics, in the form $1 - f_3$(Mbuti.DG; pop1, pop2) and performed with classical MDS algorithm (cmdscale) implemented in R 3.5.1. The hunter-gatherers were grouped based on their geographic origins and dates

(Supplementary Information, section 5). The f_3-statistics were calculated with qp3Pop 435 in ADMIXTOOLS 5.1 package[103].

The pairwise f_2 distance matrix of Epigravettian-associated groups was generated with qpfstats 200 in ADMIXTOOLS 7.0.2 package, with parameters “allsnps: YES, scale: NO”, and Mbuti.DG set as the outgroup. The neighbour-joining tree was then reconstructed using the neighbour-joining method implemented in Ape 5.3 package[104] of R 3.5.1.

The f_4-statistics were calculated by qpDstat 755 with parameter “f4 mode: YES”, with the Mbuti.DG population from Africa used as outgroup in all f-statistics analyses. The tool qpAdm 810 in ADMIXTOOLS 5.1 was applied to model the ancestries of admixed populations, with “allsnps” mode and the outgroup set selection described in Supplementary Information, section 11. Admixture graphs were reconstructed using qpGraph 6450, with allsnps mode to correct for low-coverage sample and Mbuti.DG set as the outgroup. Admixture events were dated using the ancestry covariance pattern-based DATES 753 program[105], with a bin size of 0.1 cM for covariance calculation and the start of exponential fitting at $d \geq 0.5$cM.

Phenotypic SNP analysis

As the coverage for most ancient samples was not sufficient for diploid genotype calling, we counted the reads covering selected phenotypic SNPs on reference or alternative alleles and computed the group-based allele frequencies following a maximum likelihood approach described in Mathieson et al.[37]. Details on individuals involved in the analysis, read counts processing and allele frequency computation are provided in the Supplementary Information, section 12 and Supplementary Data 1.J and 3.G.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The aligned sequences of all individuals with new genomic data reported in this study are available at the European Nucleotide Archive (ENA) under study accession number PRJEB51862. The compiled genotype file used for analyses, including re-genotyped published genomes, has been uploaded at the Edmond Data Repository of the Max Planck Society (https://edmond.mpdl.mpg.de/dataset.xhtml?persistentId=doi:10.17617/3.Y1KJMF).

Code availability

The code for the newly developed ROH based contamination estimate method is available at https://github.com/hyl317/hapROH. A user manual including an installation guide is available at https://haproh.readthedocs.io/en/latest/hapROH_with_contamination.html. The version used for this work is archived at https://edmond.mpdl.mpg.de/dataset.xhtml?persistentId=doi:10.17617/3.Y1KJMF.

References

1. Prüfer, K. et al. A genome sequence from a modern human skull over 45,000 years old from Zlatý kůň in Czechia. *Nat. Ecol. Evol.* **5**, 820–825 (2021).

2. Hajdinjak, M. et al. Initial Upper Palaeolithic humans in Europe had recent Neanderthal ancestry. *Nature* **592**, 253–257 (2021).

3. Olalde, I. & Posth, C. Latest trends in archaeogenetic research of west Eurasians. *Curr. Opin. Genet. Dev.* **62**, 36–43 (2020).

4. Fu, Q. et al. The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).

5. Green, R. E. et al. A draft sequence of the Neandertal genome. *Science* **328**, 710–722 (2010).

6. Fu, Q. et al. An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216–219 (2015).

7. Fu, Q. et al. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445–449 (2014).

8. Seguin-Orlando, A. et al. Genomic structure in Europeans dating back at least 36,200 years. *Science* **346**, 1113–1118 (2014).

9. Vallini, L. et al. Genetics and material culture support repeated expansions into Paleolithic Eurasia from a population hub out of Africa. *Genome Biol. Evol.* **14**, evac045 (2022).

10. Maier, A. & Zimmermann, A. Populations headed south? The Gravettian from a palaeodemographic point of view. *Antiquity* **91**, 573–588 (2017).

11. Dolukhanov, P. in *Cultural Transformations and Interactions in Eastern Europe* (eds Chapman, J. & Dolukhanov, P.) 122–145

(Avebury, 1993).

12. Gamble, C., Davies, W., Pettitt, P., Hazelwood, L. & Richards, M. The archaeological and genetic foundations of the European population during the Late Glacial: implications for 'agricultural thinking'. *Cambridge Archaeol. J.* **15**, 193–223 (2005).

13. Wren, C. D. & Burke, A. Habitat suitability and the genetic structure of human populations during the Last Glacial Maximum (LGM) in Western Europe. *PLoS ONE* **14**, e0217996 (2019).

14. Villalba-Mouco, V. et al. Survival of Late Pleistocene hunter-gatherer ancestry in the Iberian Peninsula. *Curr. Biol.* **29**, 1169–1177.e7 (2019).

15. Bortolini, E. et al. Early Alpine occupation backdates westward human migration in Late Glacial Europe. *Curr. Biol.* **31**, 2484–2493.e7 (2021).

16. Feldman, M. et al. Late Pleistocene human genome suggests a local origin for the first farmers of central Anatolia. *Nat. Commun.* **10**, 1218 (2019).

17. Petr, M., Pääbo, S., Kelso, J. & Vernot, B. Limits of long-term selection against Neandertal introgression. *Proc. Natl Acad. Sci. USA* **116**, 1639–1644 (2019).

18. Harris, K. & Nielsen, R. The genetic cost of Neanderthal introgression. *Genetics* **203**, 881–891 (2016).

19. Kozłowski, J. K. The origin of the Gravettian. *Quat. Int.* **359**, 3–18 (2015).

20. Goutas, N. in *Les Gravettiens* (ed. Otte, M.) 105–160 (Errance, 2013).

21. Klaric, L., Goutas, N., Laccarière, J. & Banks, W. E. in *Les Sociétés Gravettiennes du Nord-Ouest Européen: Nouveaux Sites, Nouvelles Données, Nouvelles Lectures* (eds Touzé, O., Goutas, N., Salomon, H. & Noiret, P.) 323–266 (Presses Univ. de Liège, 2021).

22. Mounier, A. et al. Gravettian cranial morphology and human group affinities during the European Upper Palaeolithic. *Sci. Rep.* **10**, 21931 (2020).

23. Sikora, M. et al. Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. *Science* **358**, 659–662 (2017).

24. Posth, C. et al. Pleistocene mitochondrial genomes suggest a single major dispersal of non-Africans and a Late Glacial Population turnover in Europe. *Curr. Biol.* **26**, 827–833 (2016).

25. Straus, L. G. The human occupation of southwestern Europe during the Last Glacial Maximum: Solutrean cultural adaptations in France and Iberia. *J. Anthropol. Res.* **71**, 465–492 (2015).

26. Lécuyer, C., Hillaire-Marcel, C., Burke, A., Julien, M. A. & Hélie, J. F. Temperature and precipitation regime in LGM human refugia of southwestern Europe inferred from $\delta^{13}C$ and $\delta^{18}O$ of large mammal remains. *Quat. Sci. Rev.* **255**, 106796 (2021).

27. Djindjian, F. Territories and economies of hunter-gatherer groups during the last glacial maximum in Europe. *Quat. Int.* **412**, 37–43 (2016).

28. Ruiz-Redondo, A. et al. Mid and Late Upper Palaeolithic in the Adriatic Basin: chronology, transitions and human adaptations to a changing landscape. *Quat. Sci. Rev.* **276**, 107319 (2022).

29. Laplace, G. *Essai de Typologie Systématique* (Annali dell'Università di Ferrara, 1964).

30. Yu, H. et al. Genomic and dietary discontinuities during the Mesolithic and Neolithic in Sicily. *iScience* **25**, 104244 (2022).

31. Palma di Cesnola, A. *Le paléolithique supérieur en Italie. Série* '*Préhistoire d'Europe*' *9* (Éditions, 2001).

32. Peresani, M. et al. Hunter-gatherers across the great Adriatic-Po region during the Last Glacial Maximum: environmental and cultural dynamics. *Quat. Int.* **581–582**, 128–163 (2021).

33. Otte, M. Appearance, expansion and dilution of the Magdalenian civilization. *Quat. Int.* **272–273**, 354–361 (2012).

34. Maier, A. in *The Central European Magdalenian* 81–180 https://doi.org/10.1007/978-94-017-7206-8_6 (Springer, 2015).

35. Kozłowski, S. K., Połtowicz-Bobak, M., Bobak, D. & Terberger, T. New information from Maszycka Cave and the Late Glacial recolonisation of Central Europe. *Quat. Int.* **272**, 288–296 (2012).

36. Raghavan, M. et al. Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87–91 (2014).

37. Mathieson, I. et al. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).

38. Mathieson, I. et al. The genomic history of southeastern Europe. *Nature* **555**, 197–203 (2018).

39. Mittnik, A. et al. The genetic prehistory of the Baltic Sea region. *Nat. Commun.* **9**, 442 (2018).

40. Günther, T. et al. Population genomics of Mesolithic Scandinavia: investigating early postglacial migration routes and high-latitude adaptation. *PLoS Biol.* **16**, e2003703 (2018).

41. Charlton, S. et al. Dual ancestries and ecologies of the Late Glacial Palaeolithic in Britain. *Nat. Ecol. Evol.* **6**, 1658–1668 (2022).

42. Damgaard, P. et al. The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* **360**, eaar7711 (2018).

43. Saag, L. et al. Genetic ancestry changes in Stone to Bronze Age transition in the East European plain. *Sci. Adv.* **7**, eabd6535 (2021).

44. Wood, R. E. et al. Freshwater radiocarbon reservoir effects at the burial ground of Minino, Northwest Russia. *Radiocarbon* **55**, 163–177 (2013).

45. Narasimhan, V. M. et al. The formation of human populations in South and Central Asia. *Science* **365**, eaat7487 (2019).

46. Gronenborn, D. in *The Spread of the Neolithic to Central Europe* (RGZM, 2010).

47. Schmitt, T. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Front. Zool.* **4**, 11 (2007).

48. Roebroeks, W., Mussi, M., Svoboda, J. & Fennema, K. *Hunters of the Golden Age: The Mid Upper Palaeolithic of Eurasia, 30,000-20,000 bp* (Univ. of Leiden, 2000).

49. Kotula, A., Piezonka, H. & Tergerger, T. The Mesolithic cemetery of Groß Fredenwalde (north-eastern Germany) and its cultural affiliations. *Liet. Archeol.* **46**, 65–84 (2020).

50. Piezonka, H. et al. The emergence of hunter-gatherer pottery in the Urals and West Siberia: new dating and stable isotope evidence. *J. Archaeol. Sci.* **116**, 105100 (2020).

51. Villalba-Mouco, V. et al. A 23,000-year-old southern-Iberian individual links human groups that lived in Western Europe before and after the Last Glacial Maximum. *Nat. Ecol. Evol.,* https://doi.org/10.1038/s41559-023-01987-0 (2023)

52. Bronk Ramsey, C. Bayesian analysis of radiocarbon dates. *Radiocarbon* **51**, 337–360 (2009).

53. Reimer, P. J. et al. The IntCal20 Northern Hemisphere radiocarbon age calibration curve (0–55 cal kbp). *Radiocarbon* **62**, 725–757 (2020).

54. Korlević, P. et al. Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. *Biotechniques* **59**, 87–

93 (2015).

55. Rohland, N. & Hofreiter, M. Ancient DNA extraction from bones and teeth. *Nat. Protoc.* **2**, 1756–1762 (2007).

56. Dabney, J. et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl Acad. Sci. USA* **110**, 15758–63 (2013).

57. Rohland, N., Glocke, I., Aximu-Petri, A. & Meyer, M. Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing. *Nat. Protoc.* **13**, 2447–2461 (2018).

58. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. *Philos. Trans. R. Soc. B* **370**, 20130624 (2015).

59. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010**, pdb.prot5448 (2010).

60. Kircher, M., Sawyer, S. & Meyer, M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* **40**, e3 (2012).

61. Gansauge, M., Aximu-Petri, A., Nagel, S. & Meyer, M. Manual and automated preparation of single-stranded DNA libraries for the sequencing of DNA from ancient biological remains and other sources of highly degraded DNA. *Nat. Protoc.* **15**, 2279–2300 (2020).

62. Meyer, M. et al. A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222–226 (2012).

63. Fu, Q. et al. DNA analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl Acad. Sci. USA* **110**, 2223–2227 (2013).

64. Peltzer, A. et al. EAGER: efficient ancient genome reconstruction. *Genome Biol.* **17**, 60 (2016).

65. Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, 88 (2016).

66. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

67. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. MapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).

68. Skoglund, P. et al. Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proc. Natl Acad. Sci. USA* **111**, 2229–2234 (2014).

69. Renaud, G., Slon, V., Duggan, A. T. & Kelso, J. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* **16**, 224 (2015).

70. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 356 (2014).

71. Huang, Y. & Ringbauer, H. hapCon: estimating contamination of ancient genomes by copying from reference haplotypes. *Bioinformatics* **38**, 3768–3777 (2022).

72. Nakatsuka, N. et al. ContamLD: estimation of ancient nuclear DNA contamination using breakdown of linkage disequilibrium. *Genome Biol.* **21**, 199 (2020).

73. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

74. Catalano, G. et al. Late Upper Palaeolithic hunter-gatherers in the Central Mediterranean: new archaeological and genetic data from the Late Epigravettian burial Oriente C (Favignana, Sicily). *Quat. Int.* **537**, 24–32 (2020).

75. Jensen, T. Z. T. et al. A 5700 year-old human genome and oral microbiome from chewed birch pitch. *Nat. Commun.* **10**, 5520–10 (2019).

76. Key, F. M. et al. Emergence of human-adapted *Salmonella enterica* is linked to the Neolithization process. *Nat. Ecol. Evol.* **4**, 324–333 (2020).

77. Rivollat, M. et al. Ancient genome-wide DNA from France highlights the complexity of interactions between Mesolithic hunter-gatherers and Neolithic farmers. *Sci. Adv.* **6**, eaaz5344 (2020).

78. Svensson, E. et al. Genome of Peştera Muierii skull shows high diversity and low mutational load in pre-glacial Europe. *Curr. Biol.* **31**, 2973–2983.e9 (2021).

79. Antonio, M. L. et al. Ancient Rome: A genetic crossroads of Europe and the Mediterranean. *Science* **366**, 708–714 (2019).

80. Brace, S. et al. Ancient genomes indicate population replacement in Early Neolithic Britain. *Nat. Ecol. Evol.* **3**, 765–771 (2019).

81. Brunel, S. et al. Ancient genomes from present-day France unveil 7,000 years of its demographic history. *Proc. Natl Acad. Sci. USA* **117**, 12791–12798 (2020).

82. Cassidy, L. M. et al. A dynastic elite in monumental Neolithic society. *Nature* **582**, 384–388 (2020).

83. González-Fortes, G. et al. Paleogenomic evidence for multi-generational mixing between Neolithic farmers and Mesolithic hunter-gatherers in the Lower Danube Basin. *Curr. Biol.* **27**, 1801–1810.e10 (2017).

84. Jones, E. R. et al. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat. Commun.* **6**, 8912 (2015).

85. Jones, E. R. et al. The Neolithic transition in the Baltic was not driven by admixture with early European farmers. *Curr. Biol.* **27**, 576–582 (2017).

86. Lazaridis, I. et al. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).

87. Lazaridis, I. et al. Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–424 (2016).

88. Lipson, M. et al. Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368–372 (2017).

89. van de Loosdrecht, M. et al. Pleistocene North African genomes link Near Eastern and sub-Saharan African human populations. *Science* **360**, 548–552 (2018).

90. Olalde, I. et al. The genomic history of the Iberian Peninsula over the past 8,000 years. *Science* **363**, 1230–1234 (2019).

91. Saag, L. et al. Extensive farming in Estonia started through a sex-biased migration from the steppe. *Curr. Biol.* **27**, 2185–2193.e6 (2017).

92. Sikora, M. et al. The population history of northeastern Siberia since the Pleistocene. *Nature* **570**, 182–188 (2019).

93. Skoglund, P. et al. Genomic diversity and admixture differs for stone-age Scandinavian foragers and farmers. *Science* **344**, 747–750 (2014).

94. Yang, M. A. et al. 40,000-year-old individual from Asia provides insight into early population structure in Eurasia. *Curr. Biol.* **27**, 3202–3208.e9 (2017).

95. Weissensteiner, H. et al. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44**, W58–W63 (2016).

96. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

97. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).

98. Rohrlach, A. B. et al. Using Y-chromosome capture enrichment to resolve haplogroup H2 shows new evidence for a two-path Neolithic expansion to Western Europe. *Sci. Rep.* **11**, 15005 (2021).

99. Martiniano, R., De Sanctis, B., Hallast, P. & Durbin, R. Placing ancient DNA sequences into reference phylogenies. *Mol. Biol. Evol.* **39**, msac017 (2022).

100. Karmin, M. et al. A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Res.* **25**, 459–466 (2015).

101. Ringbauer, H., Novembre, J. & Steinrücken, M. Parental relatedness through time revealed by runs of homozygosity in ancient DNA. *Nat. Commun.* **12**, 5425 (2021).

102. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).

103. Patterson, N. et al. Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).

104. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).

105. Chintalapati, M., Patterson, N. & Moorjani, P. The spatiotemporal patterns of major human admixture events during the European Holocene. *eLife* **11**, e77625 (2022).

Acknowledgements

The authors thank G. Marciani and O. Jöris for comments on archaeology; C. Jeong, M. Spyrou and K. Prüfer for comments on genetics; M. O'Reilly for graphical support for Fig. 5 and Extended Data Fig. 9; the entire IT and laboratory teams at the Department of Archaeogenetics of MPI-SHH for technical assistance; M. Meyer and S. Nagel for support with single-stranded library preparation; K. Post, P. van Es, J. Glimmerveen, M. Medendorp, M. Sier, S. Dikstra, M. Dikstra, R. van Eerden, D. Duineveld and A. Hoekman for providing access to human specimens from the North Sea (The Netherlands); M. D. Garralda and A. Estalrrich for providing access to human specimens from La Riera (Spain); J. Górski and M. Zając for providing access to human specimens from Maszycka cave; C. Di Patti for providing access to human specimens from San Teodoro 2 (Italy); P. Blaževičius for providing access to the Donkalnis human remains and the new radiocarbon dates; the Italian Ministry of Culture and Soprintendenza Archeologia Belle Arti e Paesaggio for the Provinces of Verona, Rovigo, and Vicenza for granting access to the human remains of Tagliente 2; F. Fontana, who carries out investigations of the Riparo Tagliente site (Italy); the Friuli Venezia Giulia Superintendency for providing access to the human tooth Pradis 1; and the Soprintendenza Archeologia Belle Arti e Paesaggio for the Provinces of Barletta-Andria-Trani and Foggia for providing access to the Paglicci human remains. This project has received funding by the European Research Council under the European Union's Horizon 2020 research and innovation programme under grant agreements no. 803147-RESOLUTION (to S.T.), no. 771234-PALEoRIDER (to W.H.), no. 864358 (to K.M.), no. 724703 and no. 101019659 (to K.H.). K.H. is also supported by the Deutsche Forschungsgemeinschaft (DFG FOR 2237). E.A. has received funding from the Van de Kamp fonds. PACEA co-authors of this research benefited from the scientific framework of the University of Bordeaux's IdEx Investments for the Future programme/GPR Human Past. A.G.-O. is supported by a Ramón y Cajal fellowship (RYC-2017-22558). L. Sineo, M.L. and D.C. have received funding from the Italian Ministry of University and Research (MUR) PRIN 2017 grants 20177PJ9XF and 20174BTC4R_002. H. Rougier received support from the College of Social and Behavioral Sciences of CSUN and the CSUN Competition for RSCA Awards. C.L.S. and T. Saupe received support from the European Union through the European Regional Development Fund (project no. 2014-2020.4.01.16-0030) and C.L.S. received support from the Estonian

Research Council grant PUT (PRG243). S. Shnaider received support from the Russian Science Foundation (no. 19-78-10053).

Funding

Open access funding provided by Max Planck Society.

Author information

Author notes

1. These authors contributed equally: Cosimo Posth, He Yu

Authors and Affiliations

1. Archaeo- and Palaeogenetics, Institute for Archaeological Sciences, Department of Geosciences, University of Tübingen, Tübingen, Germany

   Cosimo Posth

2. Senckenberg Centre for Human Evolution and Palaeoenvironment at the University of Tübingen, Tübingen, Germany

   Cosimo Posth, Dorothée G. Drucker, Hervé Bocherens, Christoph Wißing, Nicholas J. Conard & Katerina Harvati

3. Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

   Cosimo Posth, He Yu, Ayshin Ghalichi, Hélène Rougier, Yilei Huang, Harald Ringbauer, Adam B. Rohrlach, Kathrin Nägele, Vanessa Villalba-Mouco, Rita Radzeviciute, Tiago Ferraz, Alexander Stoessel, Rezeda Tukhbatova, Wolfgang Haak & Johannes Krause

4. State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing, China

He Yu

5. Department of Anthropology, California State University Northridge, Northridge, CA, USA

Hélène Rougier

6. Université de Bordeaux, CNRS, MC, PACEA UMR 5199, Pessac, France

Isabelle Crevecoeur, André Morala, Stéphane Madelaine, Priscilla Bayle, Bruno Maureille, Jean-Guillaume Bordes & Mona Le Luyer

7. School of Mathematical Sciences, University of Adelaide, Adelaide, South Australia, Australia

Adam B. Rohrlach

8. Instituto Universitario de Investigación en Ciencias Ambientales de Aragón, IUCA-Aragosaurus, Zaragoza, Spain

Vanessa Villalba-Mouco

9. Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, Germany

Rita Radzeviciute

10. Institute of Zoology and Evolutionary Research, University of Jena, Jena, Germany

Alexander Stoessel

11. Center of Excellence 'Archaeometry', Kazan Federal University, Kazan, Russia

Rezeda Tukhbatova & David Caramelli

12. Department of Biology, University of Florence, Florence, Italy

Martina Lari, Alessandra Modi & Stefania Vai

13. Estonian Biocentre, Institute of Genomics, University of Tartu, Tartu, Estonia

Tina Saupe, Christiana L. Scheib & Luca Pagani

14. St John's College, University of Cambridge, Cambridge, UK

Christiana L. Scheib

15. Department of Biological, Chemical and Pharmaceutical Sciences and Technologies, University of Palermo, Palermo, Italy

Giulio Catalano & Luca Sineo

16. Department of Biology, University of Padova, Padova, Italy

Luca Pagani

17. Department of Chemistry G. Ciamician, Alma Mater Studiorum, University of Bologna, Bologna, Italy

Sahra Talamo

18. Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

Sahra Talamo & Helen Fewlass

19. UMR 8068 CNRS, TEMPS—Technologie et Ethnologie des Mondes Préhistoriques, Nanterre Cedex, France

Laurent Klaric, Olivier Bignon-Lau, Grégory Debout, Michel Orliac, Pierre Bodu, Bénédicte Souffi & Frédérique Valentin

20. Musée National de Préhistoire, Les Eyzies de Tayac, France

André Morala & Stéphane Madelaine

21. Paléotime, Villard-de-Lans, France

Mathieu Rué

22. UMR 5140 CNRS, Archéologie des Sociétés Méditerranéennes, Université Paul-Valéry, Montpellier, France

Mathieu Rué

23. UMR 7194, Histoire Naturelle de l'Homme Préhistorique (HNHP), Département Homme et Environnement, Muséum National d'Histoire Naturelle, CNRS, UPVD, Paris, France

Laurent Crépin & Marie-Anne Julien

24. Association APRAGE (Approches pluridisciplinaires de recherche archéologique du Grand-Est), Besançon, France

Jean-Baptiste Caverne

25. Inrap GE, Metz, France

Jean-Baptiste Caverne

26. Skeletal Biology Research Centre, School of Anthropology and Conservation, University of Kent, Canterbury, UK

Emmy Bocaege

27. Dipartimento di Scienze Fisiche, della Terra e dell'Ambiente, U.R. Preistoria e Antropologia, Università degli Studi di Siena, Siena, Italy

Stefano Ricci, Francesco Boschin & Annamaria Ronchitelli

28. Accademia dei Fisiocritici, Siena, Italy

Stefano Ricci & Francesco Boschin

29. Centro Studi sul Quaternario ODV, Sansepolcro, Italy

Francesco Boschin

30. Institut National de Recherches Archéologiques, Bertrange, Luxembourg

Foni Le Brun-Ricalens

31. Department of Cultural Heritage, University of Bologna, Ravenna, Italy

Gregorio Oxilia, Eugenio Bortolini, Stefano Benazzi & Matteo Romandini

32. Human Ecology and Archaeology (HUMANE), Department of Archaeology and Anthropology, Institució Milà i Fontanals de Investigación en Humanidades, Consejo Superior de Investigaciones Científicas (IMF - CSIC), Barcelona, Spain

Eugenio Bortolini

33. UMR 7209—Archéozoologie et Archéobotanique-Sociétés, Pratiques et Environnements, Muséum National d'Histoire Naturelle, Paris, France

Antoine Zazzo

34. Dipartimento di Scienze Della Vita e Dell'Ambiente, Sezione di Neuroscienze e Antropologia, Università Degli Studi di Cagliari, Cittadella Monserrato, Cagliari, Italy

Vitale Sparacello

35. Dipartimento di Civiltà e Forme Del Sapere, Università di Pisa, Pisa, Italy

Elisabetta Starnini

36. Center for Isotope Research, Groningen University, Groningen, The Netherlands

Johannes van der Plicht

37. Inrap CIF, Croissy-Beaubourg, France

Laure Pecqueur & Bénédicte Souffi

38. UMR 7206 Éco-Anthropologie, Équipe ABBA. CNRS, MNHN, Université de Paris Cité, Musée de l’Homme, Paris, France

Laure Pecqueur & Sébastien Villotte

39. PALEVOPRIM Lab UMR 7262 CNRS-INEE, University of Poitiers, Poitiers, France

Gildas Merceron & Géraldine Garcia

40. Centre de Valorisation des Collections Scientifiques, Université de Poitiers, Mignaloux Beauvoir, France

Géraldine Garcia

41. Musées de Poitiers–Ville de Poitiers, Poitiers, France

Jean-Michel Leuvrey & Coralie Bay Garcia

42. Departamento de Geología, Facultad de Ciencia y Tecnología, Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU), Leioa, Spain

Asier Gómez-Olivencia

43. Sociedad de Ciencias Aranzadi, Donostia-San Sebastian, Spain

Asier Gómez-Olivencia

44. Centro UCM-ISCIII de Investigación sobre Evolución y Comportamiento Humanos, Madrid, Spain

Asier Gómez-Olivencia

45. Institute of Archaeology, University of Rzeszów, Rzeszów, Poland

Marta Połtowicz-Bobak

46. Foundation for Rzeszów Archaeological Centre, Rzeszów, Poland

Dariusz Bobak

47. Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA

Mona Le Luyer

48. Department of Psychiatry, Harvard Medical School, Boston, MA, USA

Mona Le Luyer

49. Groninger Instituut voor Archeologie, Groningen University, Groningen, The Netherlands

Paul Storm

50. Stralsund Museum, Stralsund, Germany

Claudia Hoffmann

51. Institute of Archaeology and Ethnology, Polish Academy of Science, Poznań, Poland

Jacek Kabaciński

52. Institute of History, Archaeology and Ethnography, Dushanbe, Tajikistan

Tatiana Filimonova

53. ArchaeoZOOlogy in Siberia and Central Asia—ZooSCAn, CNRS–IAET SB RAS International Research Laboratory, IRL 2013, Institute of Archaeology SB RAS, Novosibirsk, Russia

Svetlana Shnaider

54. Research Institute and Museum of Anthropology, Moscow State University, Moscow, Russia

Natalia Berezina & Alexandra Buzhilova

55. Grupo de I+D+i EVOADAPTA (Evolución Humana y Adaptaciones durante la Prehistoria) Departamento de Ciencias Históricas, Universidad de Cantabria, Santander, Spain

Borja González-Rabanal, Ana B. Marín-Arroyo & Lawrence G. Straus

56. Instituto Internacional de Investigaciones Prehistóricas de Cantabria (IIIPC), Universidad de Cantabria-Gobierno de Cantabria-Banco Santander, Santander, Spain

Manuel R. González Morales

57. Departamento de Biología de Organismos y Sistemas, Universidad de Oviedo, Oviedo, Spain

Belén López & Carmen Alonso-Llamazares

58. Quaternary Environments and Humans, OD Earth and History of Life, Royal Belgian Institute of Natural Sciences, Brussels, Belgium

Caroline Polet, Ivan Jadin & Sébastien Villotte

59. Musées Royaux d’Art et d’Histoire, Bruxelles, Belgium

Nicolas Cauwe

60. Institute of Historical Research, University of Girona, Catalonia, Spain

Joaquim Soler, Neus Coromina & Isaac Rufí

61. INRAP/UMR 8215 Trajectoires 21, Paris, France

Richard Cottiaux & Corinne Thevenet

62. School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, USA

Geoffrey Clark

63. Department of Anthropology, University of New Mexico, Albuquerque, NM, USA

Lawrence G. Straus

64. GéoArchPal-GéoArchÉon, Viéville sous-les-Cotes, France

Marie-Anne Julien

65. Archäologie & Münzkabinett, Universalmuseum Joanneum, Graz, Austria

Silvia Renhart

66. Museum ‘Das Dorf des Welan’, Wöllersdorf-Steinabrückl, Austria

Dorothea Talaa

67. Pradis Cave Museum, Clauzetto, Italy

Matteo Romandini

68. Department of Humanities, University of Ferrara, Ferrara, Italy

Matteo Romandini

69. National Museum of Antiquities, Leiden, The Netherlands

Luc Amkreutz

70. Faculty of Archaeology, Leiden University, Leiden, The Netherlands

Luc Amkreutz

71. Biogeology, Department of Geosciences, University of Tübingen, Tübingen, Germany

Hervé Bocherens & Christoph Wißing

72. Unité de Recherches Art, Archéologie Patrimoine, Université de Liège, Liège, Belgium

Sébastien Villotte

73. I.U. de Investigación en Arqueología y Patrimonio Histórico, University of Alicante, Sant Vicent del Raspeig, Alicante, Spain

Javier Fernández-López de Pablo & Magdalena Gómez-Puche

74. Arpa Patrimonio S. L., Alicante, Spain

Marco Aurelio Esquembre-Bebia

75. Amsterdam Centre of Ancient Studies and Archaeology, University of Amsterdam, Amsterdam, The Netherlands

Liesbeth Smits

76. Department of Anatomy, Histology and Anthropology, Faculty of Medicine, Vilnius University, Vilnius, Lithuania

Rimantas Jankauskas & Justina Kozakaitė

77. Service Régional de l’Archéologie de Bourgogne-Franche-Comté, Besançon Cedex, France

Christophe Cupillard

78. Laboratoire de Chrono-Environnement, UMR 6249 du CNRS, UFR des Sciences et Techniques, Besançon Cedex, France

Christophe Cupillard

79. Weyhe, Germany

Hartmut Benthien

80. Ulm, Germany

Kurt Wehrberger

81. LVR-LandesMuseum Bonn, Bonn, Germany

Ralf W. Schmitz & Susanne C. Feine

82. Institute of Pre- and Protohistory, University of Tübingen, Tübingen, Germany

Susanne C. Feine

83. Department of Archeological Sciences, Thuringian State Office for Monuments Preservation and Archeology, Weimar, Germany

Tim Schüler

84. University of Bucharest, Faculty of Geology and Geophysics, Department of Geology, Bucharest, Romania

Dan Grigorescu

85. Institute for Advanced Studies in Levant Culture and Civilization, Bucharest, Romania

Dan Grigorescu

86. German Archaeological Institute, Berlin, Germany

Friedrich Lüth

87. Brandenburg Authorities for Heritage Management and Archaeological State Museum, Zossen, Germany

Andreas Kotula & Franz Schopper

88. Institute for Pre- and Protohistory, Kiel University, Kiel, Germany

Henny Piezonka

89. Institute of Archeology at Brno, Czech Academy of Sciences, Centre for Palaeolithic and Paleoanthropology, Brno, Czechia

Jiří Svoboda & Sandra Sázelová

90. Institute of Archaeology, Academy of Sciences of the Republic of Tatarstan, Kazan, Russia

Andrey Chizhevsky

91. Samara State University of Social Sciences and Education, Samara, Russia

Aleksandr Khokhlov

92. Early Prehistory and Quaternary Ecology, Department of Geosciences, University of Tübingen, Tübingen, Germany

Nicholas J. Conard

93. Paleoanthropology, Institute for Archaeological Sciences, Department of Geosciences, University of Tübingen, Tübingen, Germany

Katerina Harvati

94. DFG Centre for Advanced Studies ‘Words, Bones, Genes, Tools’, University of Tübingen, Tübingen, Germany

Katerina Harvati

95. Royal Belgian Institute of Natural Sciences, Brussels, Belgium

Patrick Semal

96. Anthropologie-Büro, Berlin, Germany

Bettina Jungklaus

97. Institute of Archaeology Russian, Academy of Sciences, Moscow, Russia

Alexander Suvorov

98. School of Archaeology, University of Oxford, Oxford, UK

Rick Schulting

99. Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), Russian Academy of Sciences, Saint Petersburg, Russia

Vyacheslav Moiseyev

100. Department of Cultures, University of Helsinki, Helsinki, Finland

Kristiina Mannermaa

101. Seminar for Pre- and Protohistory, Göttingen University, Göttingen, Germany

Thomas Terberger

102. Lower Saxony State Service for Cultural Heritage, Hannover, Germany

Thomas Terberger

103. Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands

Eveline Altena

Contributions

C. Posth, H.Y., A.G., A.B.R., W.H. and J. Krause performed or supervised genetic analyses. H. Rougier and I.C. performed mortuary practices analyses. Y.H. and H. Ringbauer developed the contamination-estimation method. K.N., V.V.-M., R.R., T. Ferraz, R.T., D.G.D., M.L., A. Modi, S. Vai, T. Saupe, C.L.S., G. Catalano, L. Pagani, D.C. and E.A. performed or supervised laboratory work. H. Rougier, I.C., L.K., A. Morala, M. Rué, S.M., L.C., J.-B.C., E. Bocaege, S. Ricci, F.B., P. Bayle, B.M., F.L.B.-R., J.-G.B., G.O., E. Bortolini, O.B.-L., G.D., M.O., A.Z., V.S., E.S., L. Sineo, J.v.d.P., L. Pecqueur, G.M., G.G., J.-M.L., C.B.G., A.G.-O., M.P.-B., D.B., M.L.L., P. Storm, C.H., J. Kabaciński, T. Filimonova, S. Shnaider, N.B., B.G.-R., M.R.G.M., A.B.M.-A., B.L., C.A.-L., A.R., C. Polet, I.J., N. Cauwe, J. Soler, N. Coromina, I.R., R.C., G. Clark, L.G.S., M.-A.J., S. Rehart, D.T., S.B., M. Romandini, L.A., H. Bocherens, C.W., S. Villotte, J.F.-L.d.P., M.G.-P., M.A.E.-B., P. Bodu, L. Smits, B.S., R.J., J. Kozakaitė, C.C., H. Benthien, K.W., R.W.S., S.C.F., T. Schüler, C.T., D.G., F.L., A. Kotula, H.P., F.S., J. Svoboda, S. Sázelová, A.C., A. Khokhlov, N.J.C., F.V., K.H., P. Semal, B.J., A. Suvorov, R. Schulting, V.M., K.M., A.B. T.T., D.C. and E.A. assembled archaeological material, descriptions and interpretations. S.T. and H.F. produced a subset of the radiocarbon dates. A. Stoessel performed µCT scanning and advised on sampling strategy. C. Posth and H.Y. wrote the manuscript with input from T.T., E.A., W.H., J. Krause and all other co-authors.

Corresponding authors

Correspondence to Cosimo Posth, He Yu or Johannes Krause.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Marta Mirazón Lahr, Ludovic Orlando and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 MtDNA and Y-chromosome phylogenies.

Bold letters refer to (**a**) mtDNA and (**b**) Ychr haplogroups, whose boxes are coloured according to the legend in Extended Data Fig. 2. The labels in italic denote previously published individuals without new data generated in this study.

Extended Data Fig. 2 The distribution of mtDNA and Y-chromosome haplogroups among different hunter-gatherer groups.

The length of each coloured bar represents the fraction of individuals carrying the corresponding haplogroup (legend on the right of each panel). The number of individuals in each group is written to the right of each bar. We only plotted groups with more than two individuals and, for this reason,

individuals from the GoyetQ116-1 cluster are included here into the Fournol cluster.

Extended Data Fig. 3 F_4-statistics comparing the affinity of pre-LGM European hunter-gatherers to Goyet Q116-1 and Sunghir.

The colours correspond to the grouping of tested populations, dots refer to the f_4-values and the dark and light error bars to 1*SE and 3*SE estimated from 5 cM block jackknife, respectively. This figure shows that the Gravettian-associated individuals from western Europe (Fournol cluster) are closely related to Goyet Q116-1 while the Gravettian-associated individuals from central-eastern and southern Europe (Věstonice cluster) are closely related to the Sunghir group, representative of the Kostenki cluster. Details are provided in Data S2.B.

Extended Data Fig. 4 Admixture graph modelling of pre-34 ka hunter-gatherer lineages.

In this admixture graph, the lineage related to Zlatý kůň splits more basally than the Bacho Kiro IUP group, who contributes to Tianyuan, Ust'Ishim and Goyet Q116-1 (indicated with red lines), but not to the Sunghir group.

Extended Data Fig. 5 Summary of ROH segments detected in Eurasian hunter-gatherers.

We visualize the total amount of ROH longer than 4 cM for (**a**) pre-LGM individuals, (**b**) Epigravettian- and Magdalenian-associated individuals, (**c**) individuals carrying high proportions of Sidelkino-related ancestry, and (**d**) individuals carrying high proportions of Oberkassel-related ancestry. Colour legend is shown in (**e**). Each bar represents one individual with the ROH grouped in four length categories (grouped by colour). The inferred pattern of short ROH (4–8 cM, visualized in blue) being common is in stark contrast to most later farmer populations, where the majority of individuals have no short ROH whatsoever (see[101]), and evidences small effective population sizes across West Eurasian hunter-gatherer groups. A dashed

line of 50 cM total ROH is drawn in each panel to help comparison between panels with different y-axis scales. Details of the grouping and ROH segments are provided in Data S3.B.

Extended Data Fig. 6 West Eurasian PCA showing the genetic positioning of post-LGM hunter-gatherers.

Present-day individuals (gray dots) genotyped on the Human Origins dataset are used to define the PCA variation onto which ancient genomes (coloured symbols) are projected. The newly reported individuals with over 15,000 SNPs on the Human Origins dataset are shown in black-outlined and filled symbols, as illustrated in the legend on the right, while representative ancient genomes are shown in outlined symbols, as illustrated in the legend at the bottom of the PCA.

Extended Data Fig. 7 Admixture dates between Oberkassel/Villabruna and ANE ancestries in the oldest individuals from the Sidelkino cluster.

The triangles show the average calibrated dates of the tested groups and the dots show the estimated admixture dates with the software DATES[105] using Oberkassel (red dots) or Villabruna (blue dots) clusters as one source and ANE-related individuals as the other source population. The generation time is set to 29 years and the error bars show the SE of the admixture date estimated from jackknife resampling ($n = 22$ autosomal chromosomes). Additional details are provided in Supplementary Information, section 11 and Supplementary Table 3.

Extended Data Fig. 8 Changes of Sidelkino and Oberkassel ancestry proportions in post-14 ka hunter-gatherers.

The bivariate plots show the expansion of Oberkassel and Sidelkino ancestries through time in two European areas (longitude below and above 30 degrees). The x-axis shows the average age of each tested individual/group and the y-axis shows the proportion of Sidelkino ancestry, relative to the total hunter-gatherer ancestry (Oberkassel + Sidelkino) in

each group. The three squares highlight Baltic HG groups associated with the Comb Ceramic Culture (CCC) that show a marked increase in Sidelkino-related ancestry compared to older Baltic HG groups.

Extended Data Fig. 9 Graphical summary depicting the main genetic transformations in post-40 ka hunter-gatherers from Europe.

This figure shows (**a**) the distribution of and interaction between hunter-gatherer genetic ancestries and (**b**) a simplified schematic representation of major chrono-cultural subdivisions of the European Upper Palaeolithic (green blocks) followed by a grouped Mesolithic to Neolithic block (in gray). The x-axes report the geographic regions as divided in Fig. 1a, and the y-axes report time in thousand years before present (kBP). In panel **a**, genetic affinity between different ancestries is indicated by thick lines or shades, while admixture is indicated with arrows. In panel **b**, the colour code does not imply archaeological similarities.

Extended Data Table 1 Summary statistics for individuals with newly reported genomic data

Supplementary information

Supplementary Information

Archaeological background, method descriptions and discussions additional to the main text, including sections 1–13, Figs. 1–32, Tables 1–4 and references.

Reporting Summary

Supplementary Data 1

Metadata information of studied individuals.

Supplementary Data 2

Raw data of *f*-statistics analyses.

Supplementary Data 3

Raw data of other types of analyses.

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Posth, C., Yu, H., Ghalichi, A. *et al.* Palaeogenomics of Upper Palaeolithic to Neolithic European hunter-gatherers. *Nature* **615**, 117–126 (2023). https://doi.org/10.1038/s41586-023-05726-0

- Received: 10 June 2022
- Accepted: 12 January 2023
- Published: 01 March 2023

- Issue Date: 02 March 2023

- DOI: https://doi.org/10.1038/s41586-023-05726-0

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

A 23,000-year-old southern Iberian individual links human groups that lived in Western Europe before and after the Last Glacial Maximum

- Vanessa Villalba-Mouco
- Marieke S. van de Loosdrecht
- Wolfgang Haak

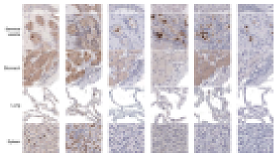
Nature Ecology & Evolution Article Open Access 01 Mar 2023

The genomic history of ice-age Europeans

- Ludovic Orlando

Nature News & Views 01 Mar 2023

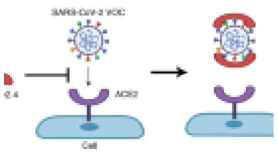
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05726-0



ACE2 localizes to the respiratory cilia and is not increased by ACE inhibitors or ARBs

28 October 2020

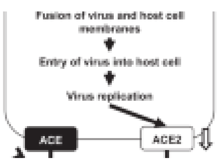
Ivan T. Lee, Tsuguhisa Nakayama, … Peter K. Jackson



Engineered ACE2 decoy mitigates lung injury and death induced by SARS-CoV-2 variants

19 January 2022

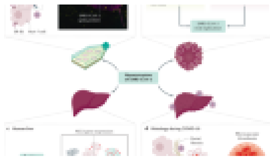
Lianghui Zhang, Soumajit Dutta, … Asrar B. Malik



Interactions of coronaviruses with ACE2, angiotensin II, and RAS inhibitors—lessons from available evidence and insights into COVID-19

27 April 2020

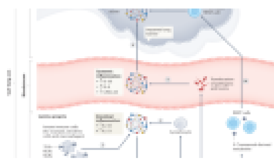
Hisashi Kai & Mamiko Kai



[COVID-19 and liver disease: mechanistic and clinical perspectives]

10 March 2021

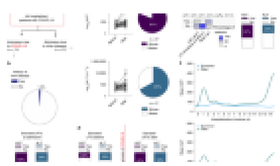
Thomas Marjot, Gwilym J. Webb, … Eleanor Barnes



[Gut microbiota in COVID-19: key microbial changes, potential mechanisms and clinical applications]

21 October 2022

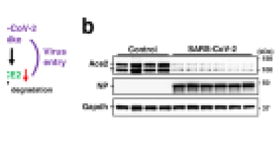
Fen Zhang, Raphaela I. Lau, … Siew C. Ng



[Molecular consequences of SARS-CoV-2 liver tropism]

28 March 2022

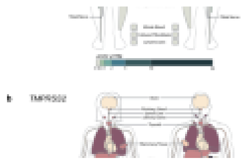
Nicola Wanner, Geoffroy Andrieux, … Tobias B. Huber



ACE2-like carboxypeptidase B38-CAP protects from SARS-CoV-2-induced lung injury

23 November 2021

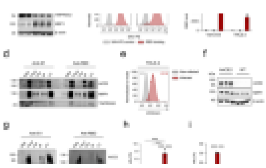
Tomokazu Yamaguchi, Midori Hoshizaki, … Keiji Kuba



From bedside to bench: regulation of host factors in SARS-CoV-2 infection

07 April 2021

Samantha Y. Q. Ong, Ilya M. Flyamer, … Simon C. Biddie



The spike of SARS-CoV-2 promotes metabolic rewiring in hepatocytes

17 August 2022

Maria Mercado-Gómez, Endika Prieto-Fernández, … María L. Martínez-Chantar

- Article
- Open Access
- Published: 05 December 2022

FXR inhibition may protect from SARS-CoV-2 infection by reducing ACE2

- Teresa Brevini ,
- Mailis Maes,
- Gwilym J. Webb,
- Binu V. John,
- Claudia D. Fuchs,
- Gustav Buescher,
- Lu Wang,
- Chelsea Griffiths,
- Marnie L. Brown,
- William E. Scott III,
- Pehuén Pereyra-Gerber,
- William T. H. Gelson,
- Stephanie Brown,
- Scott Dillon,
- Daniele Muraro,
- Jo Sharp,
- Megan Neary,
- Helen Box,
- Lee Tatham,
- James Stewart,
- Paul Curley,
- Henry Pertinez,
- Sally Forrest,
- Petra Mlcochova,
- Sagar S. Varankar,
- Mahnaz Darvish-Damavandi,
- Victoria L. Mulcahy,
- Rhoda E. Kuc,
- Thomas L. Williams,
- James A. Heslop,

- Davide Rossetti,
- Olivia C. Tysoe,
- Vasileios Galanakis,
- Marta Vila-Gonzalez,
- Thomas W. M. Crozier,
- Johannes Bargehr,
- Sanjay Sinha,
- Sara S. Upponi,
- Corrina Fear,
- Lisa Swift,
- Kourosh Saeb-Parsy,
- Susan E. Davies,
- Axel Wester,
- Hannes Hagström,
- Espen Melum,
- Darran Clements,
- Peter Humphreys,
- Jo Herriott,
- Edyta Kijak,
- Helen Cox,
- Chloe Bramwell,
- Anthony Valentijn,
- Christopher J. R. Illingworth,
- UK-PBC Consortium,
- Bassam Dahman,
- Dustin R. Bastaich,
- Raphaella D. Ferreira,
- Thomas Marjot,
- Eleanor Barnes,
- Andrew M. Moon,
- Alfred S. Barritt IV,
- Ravindra K. Gupta,
- Stephen Baker,
- Anthony P. Davenport,
- Gareth Corbett,
- Vassilis G. Gorgoulis,
- Simon J. A. Buczacki,

- Joo-Hyeon Lee,
- Nicholas J. Matheson,
- Michael Trauner,
- Andrew J. Fisher,
- Paul Gibbs,
- Andrew J. Butler,
- Christopher J. E. Watson,
- George F. Mells,
- Gordon Dougan,
- Andrew Owen,
- Ansgar W. Lohse,
- Ludovic Vallier &
- …
- Fotios Sampaziotis

Show authors

Nature **volume 615**, pages 134–142 (2023)

- 166k Accesses
- 12 Citations
- 943 Altmetric
- Metrics details

Abstract

Preventing SARS-CoV-2 infection by modulating viral host receptors, such as angiotensin-converting enzyme 2 (ACE2)[1], could represent a new chemoprophylactic approach for COVID-19 that complements vaccination[2,3]. However, the mechanisms that control the expression of ACE2 remain unclear. Here we show that the farnesoid X receptor (FXR) is a direct regulator of *ACE2* transcription in several tissues affected by COVID-19, including the gastrointestinal and respiratory systems. We then use the over-the-counter compound z-guggulsterone and the off-patent drug

ursodeoxycholic acid (UDCA) to reduce FXR signalling and downregulate ACE2 in human lung, cholangiocyte and intestinal organoids and in the corresponding tissues in mice and hamsters. We show that the UDCA-mediated downregulation of ACE2 reduces susceptibility to SARS-CoV-2 infection in vitro, in vivo and in human lungs and livers perfused ex situ. Furthermore, we reveal that UDCA reduces the expression of ACE2 in the nasal epithelium in humans. Finally, we identify a correlation between UDCA treatment and positive clinical outcomes after SARS-CoV-2 infection using retrospective registry data, and confirm these findings in an independent validation cohort of recipients of liver transplants. In conclusion, we show that FXR has a role in controlling ACE2 expression and provide evidence that modulation of this pathway could be beneficial for reducing SARS-CoV-2 infection, paving the way for future clinical trials.

Main

Since the beginning of the pandemic, the management of COVID-19 has improved considerably with the development of therapeutic agents, vaccines and monoclonal antibodies[4]. Despite the transformational effect of vaccines in populations that can access them, major global health challenges remain. New SARS-CoV-2 variants continue to emerge and are associated with high case rates and substantial global mortality. Treatment options, such as dexamethasone, remdesivir, molnupiravir and nirmatrelvir, improve clinical outcomes only in specific groups of patients[5,6]; monoclonal antibodies, such as the REGN-COV2 cocktail, show reduced neutralizing efficacy against new variants[7]; and vaccines are restricted by variable efficacy[8], the emergence of vaccine-resistant viral variants[7], cost[9] and availability[10]. Finally, one of the biggest challenges is still prophylaxis in vulnerable and high-risk groups, such as immunocompromised individuals who are not expected to mount an appropriate response to vaccines. The only prophylactic agents for these groups are monoclonal antibodies, which are hampered by the propensity of the viral spike to evolve to escape neutralization[7]. Notably, there are at present no other approved agents for pharmacological prophylaxis against COVID-19[2]. Therefore, there is a pressing need for novel prophylactic agents that reduce the risk of severe

disease[3], are less susceptible to viral resistance and are compatible with healthcare systems in low- and middle-income countries.

Viral host receptors represent logical therapeutic targets, because they are essential for SARS-CoV-2 cellular entry and infection[1]. Among these, ACE2 is particularly appealing[1]. ACE2 is a transmembrane carboxypeptidase with a broad substrate specificity, including angiotensin II, that acts as the main receptor for SARS-CoV-2. It directly binds to the spike proteins of different coronaviruses, with a high affinity for SARS-CoV-2, rendering it indispensable for viral entry[11]. Accordingly, COVID-19 predominantly affects tissues that express ACE2, such as the lungs, the cardiovascular system, the digestive tract and the biliary tree[12,13].

Modifying the expression of ACE2 could impede viral entry and protect against infection with SARS-CoV-2 and potentially other coronaviruses that use the same receptor. Furthermore, because ACE2 is a host-cell protein, its expression is not likely to be affected by mutations in the virus. Therefore, therapies that modulate ACE2 expression may be effective against multiple SARS-CoV-2 variants with a higher genetic barrier to resistance. However, the mechanisms that control ACE2 expression remain unclear. Here we use human cholangiocyte organoids as a proof-of-principle system to demonstrate that the bile acid receptor FXR controls the expression of ACE2. We show that this mechanism applies in several SARS-CoV-2-affected tissues, including gastrointestinal and respiratory epithelia. Subsequently, we demonstrate that suppressing FXR signalling, by using the approved drug UDCA or the over-the-counter phytosteroid z-guggulsterone (ZGG), reduces ACE2 expression and SARS-CoV-2 infection in vitro and in an airborne transmission model in golden Syrian hamsters. We repeat our experiments in human lungs and livers perfused ex situ and show that administering UDCA at physiologically relevant concentrations reduces ACE2 and viral infection in both organs ex vivo. We then demonstrate a reduction in the levels of ACE2 in the nasal epithelium of volunteers receiving clinically approved doses of UDCA. Finally, we interrogate an international registry cohort of patients with COVID-19 and chronic liver disease, identify a correlation between UDCA therapy and better clinical outcomes from COVID-19 and reproduce these results in a second independent cohort of liver-transplant recipients.

Bile acids modulate cholangiocyte ACE2

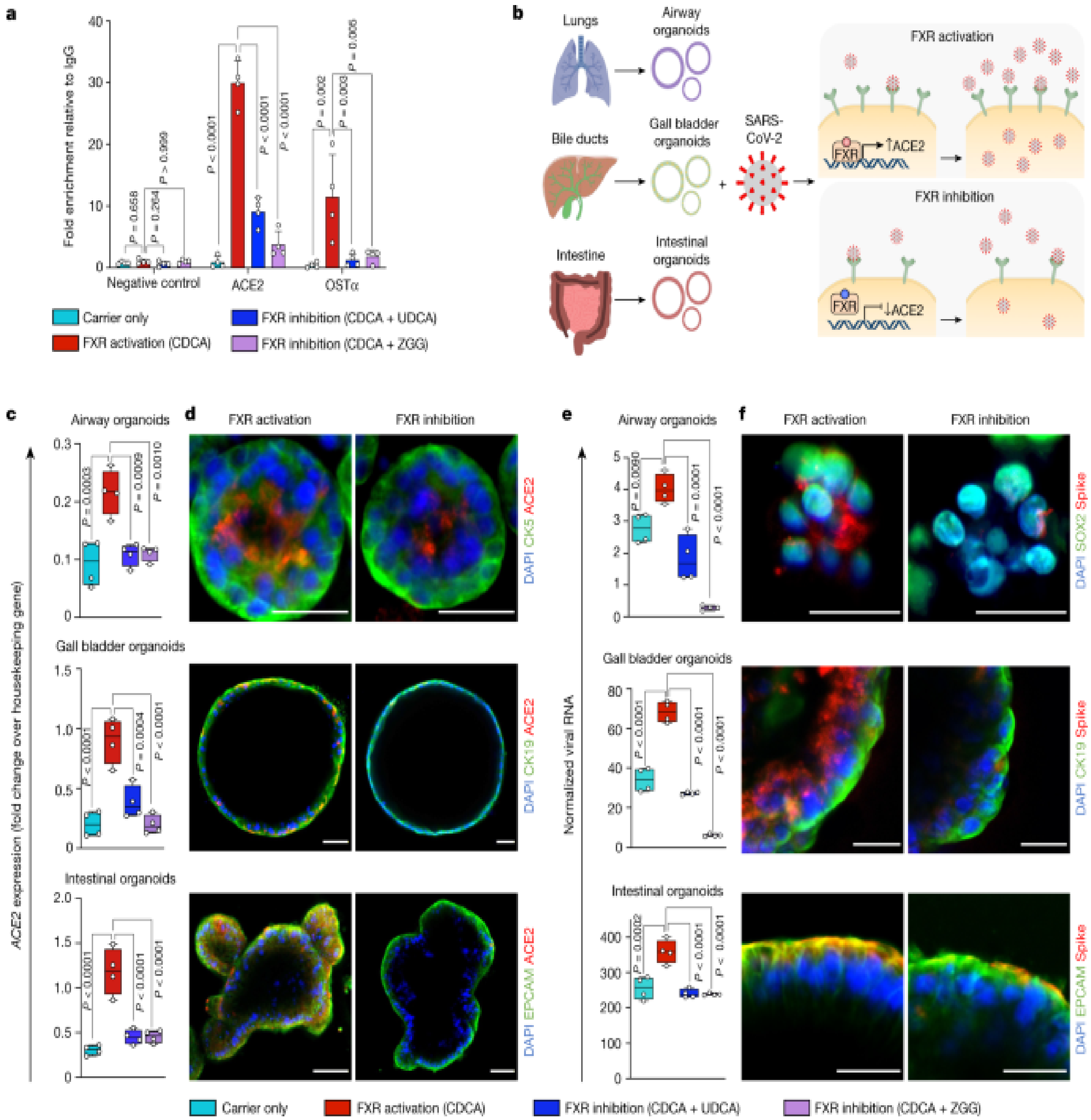
To investigate the mechanisms that control ACE2 expression, we used cholangiocyte organoids as a proof-of-principle model. Cholangiocytes are epithelial cells that line the lumen of the bile ducts and the gall bladder. We decided to focus on gall bladder cholangiocytes for several reasons. Cholangiocytes of the gall bladder express the highest levels of ACE2 in the biliary tree[12] (Wilcoxon rank-sum test $P < 4.5^{-201}$) (Extended Data Fig. 1a–d) and one of the highest levels of ACE2 in the body[12]. Thus, they can be infected by SARS-CoV-2 (Extended Data Fig. 1e,f). Furthermore, they can be propagated as organoids[14,15,16] (Extended Data Fig. 1a–c) and maintain their gall bladder identity in vitro (Extended Data Fig. 1a–c) after the bile acid chenodeoxycholic acid (CDCA) is added to their culture medium[14]. The resulting gall bladder cholangiocyte organoids (GCOs) express high levels of ACE2 (Extended Data Fig. 1c,g–i), retain their capacity to be infected by SARS-CoV-2 (Extended Data Fig. 2a–e), produce infective viral progeny (Extended Data Fig. 2c) and appropriately upregulate the expression of innate immune genes and antiviral response markers (Extended Data Fig. 2d). Of note, in the absence of bile acids (CDCA), cholangiocyte organoids lose the expression of gall bladder markers, including ACE2 (Extended Data Fig. 1g–i), which shows that CDCA is required for the expression of ACE2. These results not only show that GCOs provide an appropriate platform through which to study the mechanisms that control the expression of ACE2 in human cells, but also identify CDCA as a key regulator of SARS-CoV-2 receptor levels.

Bile acids control ACE2 levels through FXR

Because the bile acid CDCA is the most potent natural agonist of the bile acid receptor and transcription factor FXR[17], we hypothesized that CDCA could act through FXR to control the expression of ACE2. To test this hypothesis, we first confirmed that FXR is expressed in gall bladder cholangiocytes in vivo and in the corresponding GCOs in vitro (Extended Data Fig. 3a–c) and that it is activated by treatment with CDCA, as evidenced by the upregulation of its downstream target small heterodimer partner (SHP) (Extended Data Fig. 3c). To confirm that FXR is essential for

the CDCA-induced upregulation of ACE2, we knocked down the expression of FXR in cholangiocyte organoids by using short hairpin RNAs (shRNAs) and found that this prevented the upregulation of ACE2 and SHP after treatment with CDCA (Extended Data Fig. 4a–c). To assess whether FXR could bind to *ACE2* and potentially control its transcriptional activity, we analysed the *ACE2* promoter region and identified the presence of the FXR response element (FXRE) IR-1. Accordingly, we used chromatin immunoprecipitation followed by quantitative PCR (ChIP–qPCR) to confirm that activated FXR directly binds to the *ACE2* promoter (Fig. 1a), and showed the functional relevance of this binding using a luciferase reporter that contains the IR-1 region of *ACE2* (Extended Data Fig. 4e,f). Site-specific mutagenesis on the IR-1 region reduced the luciferase signal (Extended Data Fig. 4e,f), demonstrating the specificity of the FXR-binding site on the *ACE2* promoter. Conversely, suppression of FXR signalling, using the FXR antagonist ZGG[18] or the clinically used drug UDCA[17], reduced the activity of FXR (as evidenced by decreased levels of SHP; Extended Data Fig. 5a), reduced the presence of FXR on the *ACE2* promoter (Fig. 1a and Extended Data Fig. 4e,f) and downregulated the expression of ACE2 at the transcript and protein levels (Fig. 1c,d and Extended Data Fig. 5b–d). Considered together, these results show that FXR directly controls the expression of ACE2 in cholangiocytes (Fig. 1b).

Fig. 1: FXR modulates ACE2 expression and SARS-CoV-2 infection.



a, ChIP–qPCR on cholangiocyte organoids, showing that the FXR agonist CDCA promotes the binding of FXR on the *ACE2* promoter, and that this is reduced by FXR inhibitors (UDCA and ZGG). OSTα as positive control; *ACE2* promoter adjoining region as negative control; $n = 4$ independent experiments; one-way ANOVA adjusted for multiple comparisons; bars,s.d. **b**, Schematic representation of the suggested mechanism for FXR-mediated control of ACE2 expression and SARS-CoV-2 infection relative to **e**,**f**. **c**,**d**, qPCR (**c**) and immunofluorescence (**d**) showing the levels of *ACE2* after modulation of FXR activity in primary airway, biliary and intestinal organoids. Housekeeping gene, *HMBS* (also known as *PBGD*); $n = 4$

independent experiments; one-way ANOVA; centre line, median; box, interquartile range (IQR); whiskers, range; bars, s.d. Yellow scale bars, 50 µm; grey scale bars, 25 µm. **e**, qPCR quantifying SARS-CoV-2 viral RNA 24 h after infection in primary organoids treated with physiological levels of bile acids (CDCA) in the presence or absence of FXR inhibitors (UDCA and ZGG). Housekeeping gene, *GAPDH*; $n = 4$ independent experiments; one-way ANOVA adjusted for multiple comparisons; centre line, median; box, interquartile range (IQR); whiskers, range; bars, s.d. **f**, Immunofluorescence images showing the presence of SARS-CoV-2 spike protein 24 hours after infection in organoids corresponding to **e**. Scale bars, 25 µm. CDCA, UDCA and ZGG concentration, 10 µM.

FXR regulates ACE2 in various cell types

FXR is expressed in several cell types[17,19,20,21] and can be activated by bile acids, which are present not only in the gastrointestinal tract[22] but also in the lungs[20,21] and in the systemic circulation[22]. Thus, ACE2 regulation through FXR may represent a general mechanism, extending beyond cholangiocytes. To examine this possibility, we repeated our experiments using primary organoids from key organs infected by SARS-CoV-2[23], such as the lungs and the intestine. The relevance of these platforms for studying SARS-CoV-2 infection has already been shown[24,25]. We first confirmed the expression of FXR in these tissues, both in vivo and in vitro (Extended Data Fig. 3a–c). Subsequently, we showed that treatment with physiological concentrations of CDCA (10 µM)[22] resulted in the activation of FXR, as evidenced by the upregulation of the FXR downstream target SHP (Extended Data Fig. 3c) and by increased ACE2 expression (Fig. 1c,d and Extended Data Fig. 5a–c). Conversely, suppression of FXR signalling by UDCA or ZGG reduced the levels of ACE2 and SHP in primary airway and intestinal organoids (Fig. 1c,d and Extended Data Fig. 5a–d). Notably, CDCA, UDCA and ZGG had no cytotoxic effects in the concentration range that was used for our experiments (Extended Data Fig. 5e,f). These results confirm that FXR participates in the regulation of ACE2 expression in organoids derived from the respiratory, biliary and intestinal epithelium, suggesting that FXR-mediated control of ACE2 expression may be relevant for several organs (Fig. 1b).

FXR regulates viral infection in vitro

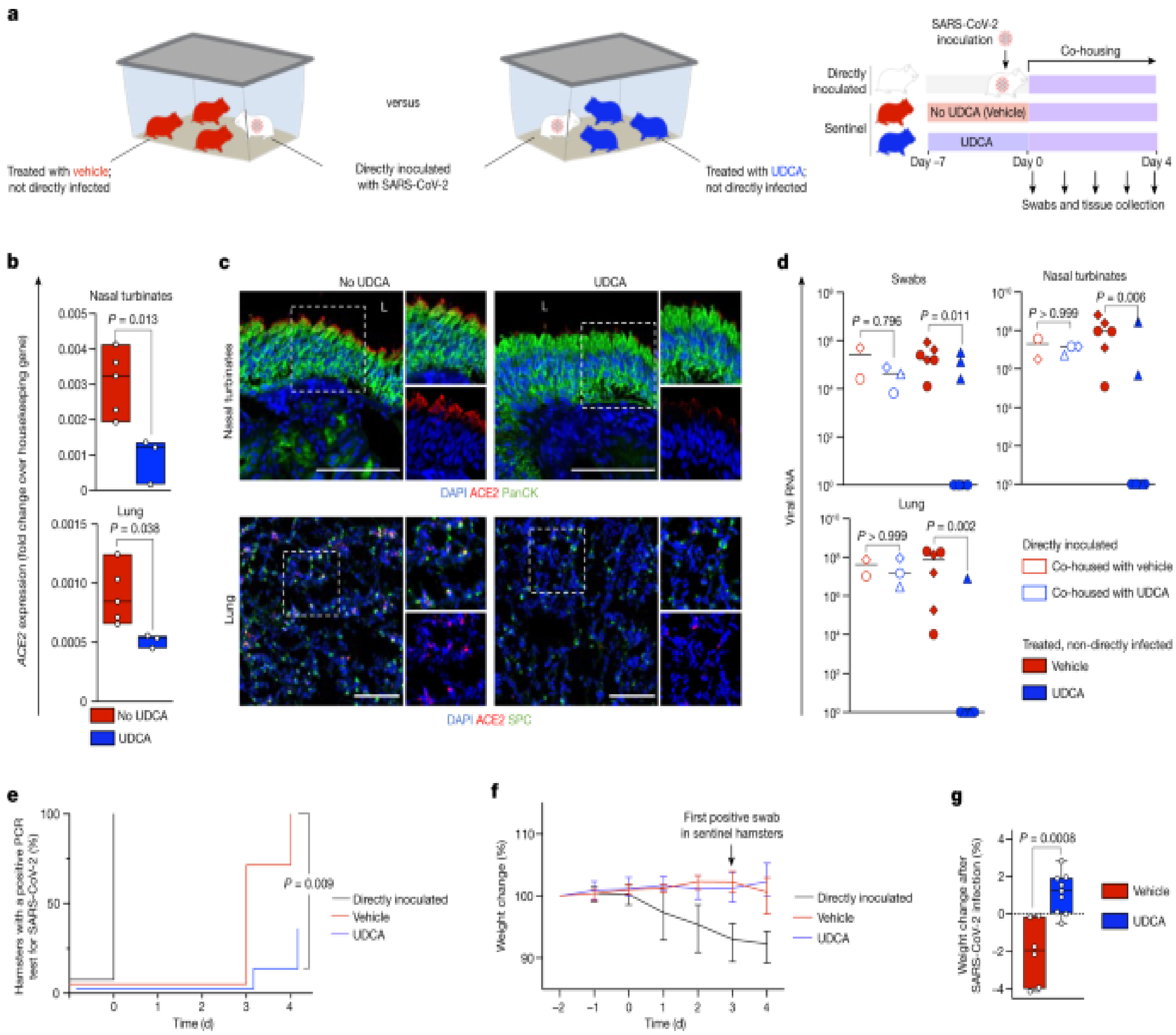
Our results show that suppressing FXR signalling with the clinically approved drug UDCA—which is used as a first-line treatment in primary biliary cholangitis (PBC)[26]—or with the over-the-counter drug ZGG reduces the expression of ACE2 in multiple cell types. To consider the relevance of this finding for COVID-19, we investigated whether the FXR-mediated downregulation of ACE2 could reduce susceptibility to SARS-CoV-2 infection in vitro. For this, we exposed gall bladder cholangiocyte, airway and intestinal organoids to physiological levels of CDCA, to simulate the baseline level of FXR activation that is present in vivo, and infected them with SARS-CoV-2 isolated from a patient's nasopharyngeal swab[27] in the absence or presence of UDCA or ZGG (Fig. 1e,f). Suppressing FXR signalling with UDCA or ZGG reduced viral infection in all three types of organoid (Fig. 1e,f and Extended Data Fig. 5a). We then investigated whether the observed reduction in viral infection was a direct result of the FXR-mediated downregulation of ACE2. First, we showed that knockdown of FXR using shRNAs decreases the expression of ACE2 and inhibits viral infection in cholangiocyte organoids independently of the presence of CDCA or that of UDCA or ZGG (Extended Data Fig. 4d). Accordingly, after knockdown, treatment with UDCA or ZGG had no effect on viral infection (Extended Data Fig. 4d). Next, to determine whether the modulation of ACE2 is the only mechanism by which UDCA and ZGG reduce SARS-CoV-2 infection, we treated HEK293T cells that had been genetically engineered to overexpress ACE2 independent of FXR[28] (Extended Data Fig. 6a,b) with UDCA or ZGG, and then infected them with SARS-CoV-2. As expected, in the absence of ACE2 modulation, UDCA and ZGG did not affect viral replication (Extended Data Fig. 6c). Together, these results confirm that UDCA and ZGG reduce susceptibility to SARS-CoV-2 infection in multiple cell types in vitro through the FXR-mediated regulation of ACE2.

FXR regulates viral infection in vivo

To validate the relevance of these findings in vivo, we assessed the effect of UDCA on ACE2 expression in FVB/N mice and Syrian golden hamsters.

We compared the expression of ACE2 in the respiratory, biliary and intestinal epithelia of four mice that were treated with UDCA relative to four control mice that did not receive UDCA (Extended Data Fig. 7a). We repeated the same experiment in Syrian golden hamsters receiving UDCA ($n = 3$ hamsters) and control untreated hamsters ($n = 5$ hamsters) and measured the levels of ACE2 in the nasal, respiratory, biliary and intestinal epithelium of these animals (Fig. 2a). Our results show that treatment with UDCA reduces the expression of ACE2 in mice (Extended Data Fig. 7a–c) and hamsters (Fig. 2a–c and Extended Data Fig. 7d,e).

Fig. 2: Inhibition of FXR reduces ACE2 expression and SARS-CoV-2 infection in vivo.



a, Schematic of the experiment performed in Syrian golden hamsters. Sentinel hamsters were not directly inoculated with virus. SARS-CoV-2

infection in sentinel hamsters was achieved through transmission from directly inoculated hamsters after co-housing. **b**, qPCR showing that treatment with UDCA reduces the levels of ACE2 in hamster nasal turbinates and lungs. Housekeeping gene, *Gapdh*; $n = 5$ vehicle (no UDCA) group versus $n = 3$ UDCA group; unpaired two-tailed t-test; centre line, median; box, interquartile range; whiskers, range; bars, s.d. **c**, Immunofluorescence images showing the levels of ACE2 in nasal and respiratory epithelium of hamsters receiving UDCA versus vehicle. $n = 3$ hamsters per group. Scale bars, 100 μm. **d**, qPCR showing the levels of SARS-CoV-2 RNA in swabs, nasal turbinates and lungs of directly inoculated hamsters and sentinel hamsters treated with UDCA or vehicle and co-housed with infected hamsters. Samples were collected after four days of co-housing. SARS-CoV-2 nucleocapsid RNA quantification relative to 18s rRNA. $n = 3$ hamsters per group; $n = 9$ UDCA, $n = 6$ vehicle hamsters; hamsters from each experiment are represented with different symbols; Kruskal–Wallis test adjusted for multiple comparisons. **e**, Kaplan–Meier curve showing the percentage of hamsters with a PCR-positive swab for SARS-CoV-2 over the course of the experiment outlined in **a**. $n = 9$ UDCA, $n = 6$ vehicle, $n = 5$ directly inoculated hamsters; log-rank Mantel–Cox test comparing UDCA versus vehicle. **f**, Percentage weight change from the start of the experiment outlined in **a**. Bars, range. Day 0 corresponds to the start of co-housing. **g**, Percentage weight change after SARS-CoV-2 infection in sentinel hamsters. The time of infection was defined as the earliest day on which a sentinel hamster had a positive swab (day 3 for both UDCA and vehicle groups). $n = 3$ independent experiments; $n = 9$ UDCA, $n = 6$ vehicle, $n = 5$ directly inoculated hamsters; unpaired two-tailed t-test; centre line, median; box, interquartile range; whiskers, range; bars, s.d.

Source data

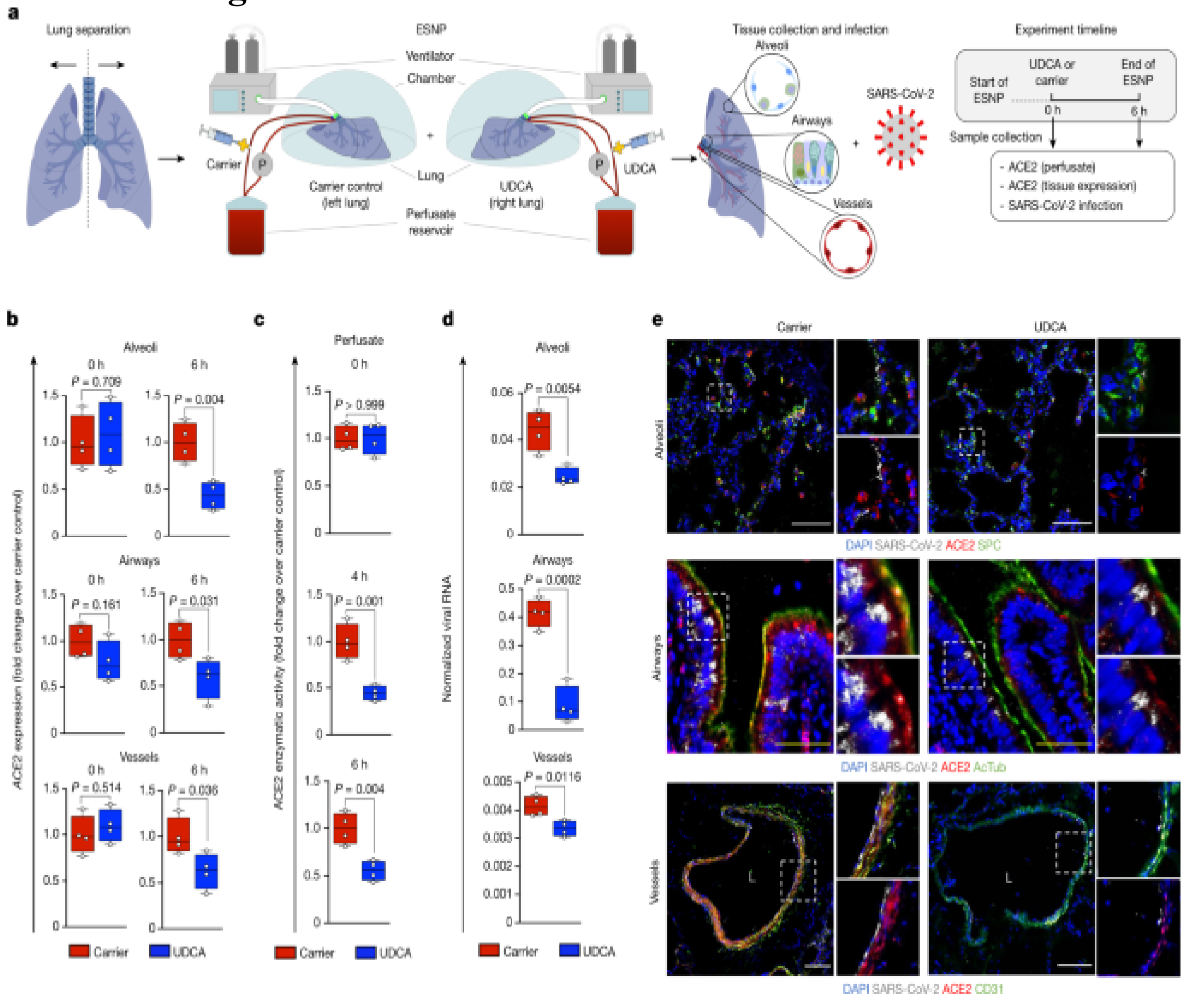
To examine whether the UDCA-mediated downregulation of ACE2 reduces SARS-CoV-2 infection in vivo, we used the well-established Syrian golden hamster model of infection. Nine hamsters were treated with UDCA for seven days (UDCA group), to achieve plasma concentrations of UDCA comparable to those in the blood of patients[29] (Extended Data Fig. 8a). Another six hamsters that received only the vehicle were used as controls (control group). Neither group was directly infected with the virus (sentinel animals). On day 7 we inoculated $n = 5$ independent, healthy hamsters with

the SARS-CoV-2 Delta variant (B.1.617.2) via the intranasal route (directly infected animals). Subsequently, each infected hamster was co-housed with a group of $n = 3$ randomly selected sentinel (uninfected) hamsters from the UDCA or the control group for a period of four days to assess SARS-CoV-2 transmission (Fig. 2a). Viral infection in sentinel hamsters was assessed with plaque assays from the lungs collected at the end of the experiment (Extended Data Fig. 8b) and confirmed with daily swabs, and viral qPCR in tissue collected from the lungs and nasal turbinates of the hamsters at the end of the experiment. Our data show that treatment with UDCA prevented the transmission of SARS-CoV-2 in $n = 6$ out of 9 sentinel hamsters (33% infected versus 67% uninfected), whereas SARS-CoV-2 was transmitted in $n = 6$ out of 6 (100%) sentinel hamsters receiving vehicle ($P = 0.027$, Fisher's exact test) (Fig. 2d,e) for the duration of the experiment. Both directly inoculated and control hamsters lost weight after viral infection, in contrast to UDCA-treated hamsters, which gained weight (Fig. 2f,g), suggesting a milder course of clinical disease in UDCA-treated hamsters. In summary, our in vivo results confirm the chemoprophylactic potential of UDCA against COVID-19.

FXR regulates infection in human organs

We then looked to validate these observations in whole human organs. We focused initially on the lung as one of the primary sites of SARS-CoV-2 infection. To conduct our experiments, we used a pair of human lungs that was declined for transplantation and performed ex situ normothermic perfusion (ESNP) with clinically appropriate mechanical ventilation to oxygenate the lungs ex vivo (Fig. 3a). ESNP was developed to objectively assess and potentially improve donor organ function, enhance organ preservation and reduce reperfusion injury by perfusing grafts with warm oxygenated blood (packed red cells) or substitute perfusion solution before transplantation[30,31]. This setting ensured that the lungs remained in near-physiological conditions during the experiment[32]. To assess the effect of UDCA, we surgically divided the right and the left lungs from the same donor and connected them to two separate but identical ESNP circuits. This setting allowed us to administer UDCA to one lung (UDCA lung) and use the other lung as a matched control treated with carrier without UDCA (control lung) to facilitate comparison (Fig. 3a).

Fig. 3: FXR inhibition reduces ACE2 levels and SARS-CoV-2 infection in a human lung ex vivo.



a, Schematic representation of the lung ESNP experiment, including type of samples and timeline. 0 h: baseline sample collection and UDCA or carrier administration. The 0-h samples were collected before administration of UDCA. 6 h: 6 h after UDCA or carrier administration. For each time point, four independent tissue samples were obtained from the lung parenchyma (alveoli), the airways and the vessels for each lung and used for ACE2 measurement and viral infection ($n = 4$ lung parenchyma, $n = 4$ airway and $n = 4$ pulmonary vessel samples per lung per time point). **b**, qPCR showing that treatment with UDCA reduces the levels of ACE2 in human alveoli, airway and pulmonary vessels perfused ex situ. Housekeeping gene, *GAPDH*; $n = 4$ independent samples; unpaired two-tailed t-test; centre line, median; box, interquartile range (IQR); whiskers, range; error bars, s.d. **c**, ACE2 enzymatic activity in the perfusate, showing that UDCA reduces

ACE2. $n = 4$ independent samples; unpaired two-tailed *t*-test; centre line, median; box, interquartile range (IQR); whiskers, range; bars, s.d. **d**, qPCR showing that 6 h of ESNP with UDCA reduces SARS-CoV-2 infection in human alveoli, airway and pulmonary vessels ex vivo. Housekeeping gene, *GAPDH*. $n = 4$ independent samples; unpaired two-tailed *t*-test. **e**, Immunofluorescence staining for ACE2 and SARS-CoV-2 in human alveoli, airway and pulmonary vessels after ESNP (6 h) with UDCA or carrier. $n = 4$ independent samples. White scale bars, 100 µm, yellow scale bars, 50 µm. UDCA concentration, 2,000 ng ml^{-1}. AcTub, acetylated α-tubulin; L, lumen.

Immediately before the administration of UDCA, we measured baseline ACE2 expression in both lungs (0-h samples collected from lung parenchyma, airway and pulmonary vessels; $n = 4$ independent samples from each part of the organ per lung, 24 samples in total; Fig. 3b) and ACE2 activity in the circulating perfusate from each ESNP circuit ($n = 4$ independent measurements per circuit; Fig. 3c). We then administered UDCA 'systemically' in the perfusate of the UDCA lung; UDCA was diluted in saline to 2,000 ng ml^{-1}, corresponding to the steady-state plasma concentration achieved in patients after multiple doses of oral UDCA[29]. At the same time, the control lung received an equal volume of saline (carrier) (Fig. 3a; 0 h refers to the administration of UDCA or carrier). We continued ex situ perfusion with UDCA for 6 h. Repeat perfusate and tissue samples were collected at 6 h, matching our pre-UDCA measurements ($n = 24$ independent tissue samples and $n = 8$ independent perfusate samples per time point). We observed that ex vivo treatment with UDCA reduced the expression of ACE2 in lung parenchyma, airway and pulmonary vessels, and the activity of ACE2 in the perfusate, compared with the carrier control (Fig. 3b,c).

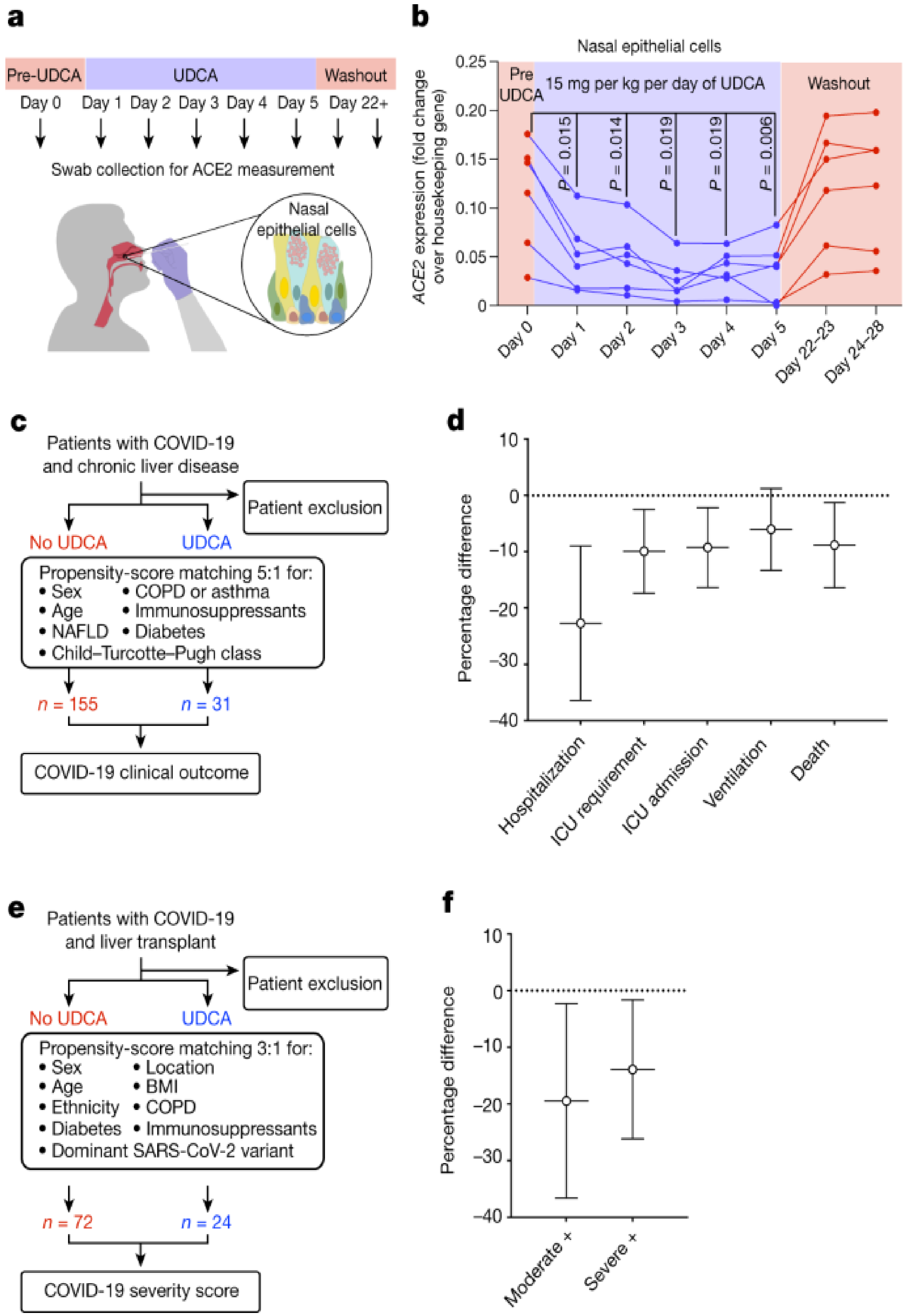
We then assessed the importance of this reduction in ACE2 for susceptibility to SARS-CoV-2 infection. We infected samples from the lung parenchyma, airway and vessels of each lung 6 h after treatment with UDCA or carrier ($n = 4$ parenchymal, $n = 4$ bronchial and $n = 4$ pulmonary vessel independent samples per lung; see Methods) and observed that treatment with UDCA reduced SARS-CoV-2 infection (Fig. 3d,e). These results show that clinical doses of circulating UDCA can downregulate the levels of ACE2 and reduce SARS-CoV-2 infection in human lungs ex vivo.

FXR inhibitors are metabolized by the liver, and ultimately distributed to different tissues through the systemic circulation. To simulate this process, we repeated ESNP with two human liver grafts (Extended Data Fig. 9a; see Methods). One liver was perfused with UDCA (2,000 ng ml^{-1}); the other was perfused with carrier and served as a control. In keeping with our lung findings, we observed that ‘systemic’ treatment with UDCA decreased the levels of ACE2 in the circulating perfusate (Extended Data Fig. 9b) and in gall bladder cholangiocytes ($n = 4$ independent samples from the grafts’ gall bladder per time point) (Extended Data Fig. 9c,d) and reduced SARS-CoV-2 infection in gall bladder cholangiocytes (Extended Data Fig. 9e,f). These results confirm that suppression of FXR signalling through systemic administration of the approved drug UDCA can downregulate ACE2 in the circulating perfusate and tissue (lung parenchyma, bronchi, vessels and gall bladder epithelium) of machine-perfused organs and reduce SARS-CoV-2 infection ex vivo.

UDCA reduces ACE2 in humans

Our previous results encouraged us to assess the potential effect of UDCA on the levels of ACE2 in humans. Given the favourable safety profile, lack of side effects and limited cost of UDCA, we recruited eight volunteers from the University Medical Centre Hamburg-Eppendorf and treated them with UDCA at the standard therapeutic dosage of 15 mg per kg per day[26] for 5 days (Supplementary Table 5). The nasal epithelial cells of the volunteers were collected using nasopharyngeal swabs, and the levels of ACE2 were measured at multiple time points before, during and after treatment with UDCA (see Methods and Fig. 4a). Participants with non-detectable cellular RNA in their nasopharyngeal swabs were excluded ($n = 2$). Our results showed that in humans, UDCA reduces the levels of ACE2 in the nasal epithelium, which is a prime site of SARS-CoV-2 infection (Fig. 4b).

Fig. 4: UDCA is associated with lower levels of ACE2 and a better clinical outcome in patients with COVID-19.



a
Pre-UDCA
UDCA
Washout
Day 0
Day 1
Day 2
Day 3
Day 4
Day 5
Day 22+
Swab collection for ACE2 measurement
Nasal epithelial cells
b
Nasal epithelial cells
Pre UDCA
15 mg per kg per day of UDCA
Washout
P = 0.015
P = 0.014
P = 0.019
P = 0.019
P = 0.006
ACE2 expression (fold change over housekeeping gene)
0.25
0.20
0.15
0.10
0.05
0
Day 0
Day 1
Day 2
Day 3
Day 4
Day 5
Day 22–23
Day 24–28
c
Patients with COVID-19 and chronic liver disease
Patient exclusion
No UDCA
UDCA
Propensity-score matching 5:1 for:
• Sex • COPD or asthma
• Age • Immunosuppressants
• NAFLD • Diabetes
• Child–Turcotte–Pugh class
n = 155
n = 31
COVID-19 clinical outcome
d
Percentage difference
10
0
−10
−20
−30
−40
Hospitalization
ICU requirement
ICU admission
Ventilation
Death
e
Patients with COVID-19 and liver transplant
Patient exclusion
No UDCA
UDCA
Propensity-score matching 3:1 for:
• Sex • Location
• Age • BMI
• Ethnicity • COPD
• Diabetes • Immunosuppressants
• Dominant SARS-CoV-2 variant
n = 72
n = 24
COVID-19 severity score
f
Percentage difference
10
0
−10
−20
−30
−40
Moderate +
Severe +

a,**b**, Schematic representation of the study design. Six healthy individuals received 15 mg per kg per day of UDCA for 5 days. ACE2 levels were measured by qPCR in nasal epithelial cells collected with nasopharyngeal swabs. Day 0 corresponds to samples collected immediately before starting UDCA treatment. Samples were collected daily during drug administration and again at day 22–23 and 24–28 to assess the washout of UDCA. **b**, qPCR measurement of the levels of *ACE2* in nasal epithelial cells collected with nasopharyngeal swabs. Each dot represents one individual measurement; lines connect dots from the same individual ($n = 6$). Housekeeping gene, *GAPDH*; $n = 6$ individuals; one-way ANOVA with Geisser–Greenhouse correction. See Supplementary Table 5 for participant characteristics. **c**, Schematic overview of the analysis performed in the exploratory cohort corresponding to **d** (see Methods). **d**, Propensity-score-matched analyses showing major outcomes after SARS-CoV-2 infection in patients taking UDCA compared to control individuals not taking UDCA. $n = 155$ patients not on UDCA; $n = 31$ patients on UDCA. See Supplementary Tables 6 and 7 for patient characteristics. Bars, 95% CI. ICU, intensive care unit. **e**, Schematic overview of the analysis performed in the validation cohort corresponding to **f** (see Methods). **f**, Propensity-score-matched analyses showing disease severity after SARS-CoV-2 infection in patients taking UDCA compared to control individuals not taking UDCA, using the NIH COVID-19 severity score. Moderate +, moderate, severe or critical disease; severe +, severe or critical disease. $n = 72$ patients not on UDCA; $n = 24$ patients on UDCA. See Supplementary Table 8 for patient characteristics. Bars, 95% CI.

To further confirm our findings, we took advantage of the fact that UDCA is extensively used in patients with cholestatic liver disorders (for example, as a first-line treatment for the cholestatic autoimmune disorder PBC). We interrogated a published serum proteomics dataset from the UK-PBC patient cohort[33], comparing the levels of ACE2 in the serum of patients who had not been treated with UDCA ($n = 62$) and patients who had received UDCA ($n = 308$). We observed that UDCA correlates with lower serum levels of ACE2 after linear regression for age, sex, body mass index (BMI), stage of liver disease (Child–Turcotte–Pugh class) and alkaline phosphatase (ALP) ($P = 0.007$; Extended Data Fig. 10a,b; see Methods), validating our previous findings (Supplementary Table 5 and Supplementary Information).

UDCA may improve COVID-19 outcome

On the basis of these observations, we decided to investigate the possible effect of UDCA treatment on the outcome of COVID-19 in patients. For this, we analysed the COVID-Hep and SECURE-Liver registries[34,35]. These registries comprise data from patients with chronic liver disease ($n = 1,096$) who developed COVID-19, including patients with cholestatic liver disorders who had been treated with UDCA ($n = 31$) (Fig. 4c and Supplementary Tables 7 and 8). We observed that—accepting the potential for selection bias in case reporting—patients who were treated with UDCA had better outcomes compared to patients who did not receive UDCA, including reduced hospitalization, admission to intensive care units and death (Fig. 4d), after propensity-score matching (no UDCA:UDCA = 5:1) for sex, age, diabetes, stage of liver disease (Child–Turcotte–Pugh class), immunosuppression, chronic pulmonary disease and non-alcoholic fatty liver disease (Fig. 4d). We note that propensity-score matching was not possible for alcohol-related liver disease (ARLD), so these patients were excluded from the analysis (see Methods). We then sought to replicate these results in a second independent patient cohort. For this, we interrogated liver-transplant recipients in the Veterans Outcomes and Costs Associated with Liver disease (VOCAL) cohort who received at least two doses of a COVID-19 mRNA vaccine. Of 119 vaccinated participants who developed COVID-19, 24 were receiving UDCA (Supplementary Table 9). These 24 participants on UDCA were matched with 72 who were not on UDCA (no UDCA:UDCA = 3:1) for sex, age, ethnicity, BMI, location within the United States, diabetes, chronic pulmonary disease, the type of immunosuppression (calcineurin inhibitor therapy, with or without anti-metabolite therapy) and the dominant SARS-CoV-2 variant at the time of infection (Fig. 4e). We found that, again accepting the potential for selection bias in case reporting, patients on UDCA were less likely to develop moderate, severe or critical COVID-19 ($P = 0.026$; Fig. 4f) according to the National Institutes of Health (NIH) COVID-19 severity score[36]. During the publication of this manuscript, we became aware of an independent study investigating the association between exposure to UDCA and outcomes of COVID-19 among participants in the VOCAL cohort with cirrhosis[37]. In this analysis, 1,607 participants with cirrhosis and UDCA exposure were propensity-score-matched with 1,607 participants with cirrhosis but without UDCA exposure.

The authors found, using multivariable logistic regression, that UDCA exposure was associated with 46% reduced odds of developing COVID-19 (adjusted odds ratio (aOR) 0.54, 95% confidence interval (CI) 0.41–0.71, $P < 0.0001$). The association was observed across the spectrum of COVID-19 against symptomatic illness (aOR 0.54, 95% CI 0.39–0.73, $P < 0.0001$), at least moderate COVID-19 (aOR 0.51, 95% CI 0.32–0.81, $P = 0.005$) and severe or critical COVID-19 (aOR 0.48, 95% CI 0.25–0.94, $P = 0.03$). These results provide additional independent evidence that reinforces the results of our study. Together, the findings from our exploratory and validation cohorts suggest that the effects of UDCA on the clinical outcomes of COVID-19 should be further investigated in a large prospective clinical trial.

Discussion

Considered collectively, our results show that FXR participates in the regulation of ACE2 expression in several tissues that are involved in SARS-CoV-2 replication. Suppressing FXR activity, using the clinically approved drug UDCA, downregulates the expression of ACE2 and reduces SARS-CoV-2 infection in vitro, in vivo and ex vivo. Furthermore, our clinical observations indicate that UDCA reduces the levels of ACE2 in the nasal epithelium of healthy individuals and suggest a correlation between UDCA and positive clinical outcomes in patients with COVID-19.

The finding that FXR regulates ACE2 is not entirely surprising. The functions of ACE2 as a molecular chaperone for the amino acid transporter SLC6A19 and as a peptidase justify its presence in the gastrointestinal tract and suggest that it has a role in digestion. Accordingly, the upregulation of ACE2 expression by FXR—which is activated by bile, a digestive fluid—may reflect a mechanism to increase peptidase levels and amino acid absorption during digestion. Moreover, in addition to its role in the gastrointestinal system, FXR is expressed in multiple organs, including the lungs[20,21], with a broad variety of functions, ranging from bile acid[22] and lipid metabolism[38] to glucose homeostasis[39], fibrosis[40] and inflammation[41]. Notably, its natural ligands, such as bile acids and hormones[42] (for example, androgens), are present in the systemic circulation[22] and it is the therapeutic target of several approved drugs[17]. This broad expression and function

explains how FXR could regulate ACE2 in tissues beyond those of the biliary tree.

Our results reveal the potential of ACE2 modulation as a host-directed treatment that might be efficacious as primary and secondary prophylaxis in COVID-19. These findings are in keeping with existing studies that show the benefits of targeting the virus–host interaction for SARS-CoV-2 at the level of ACE2 (ref. [1]) or the spike protein[7]. Indeed, large Mendelian randomization analyses in more than 7,554 patients admitted to hospital with COVID-19 and more than one million control individuals found that higher levels of ACE2 were strongly correlated with an increased risk of COVID-19-related hospitalization, identifying ACE2 as a logical candidate for drug development in COVID-19 (ref. [1]). In addition, the extensive clinical literature that supports Ronapreve and Evusheld—both dual combinations of monoclonal antibodies against the SARS-CoV-2 spike protein—shows the utility of inhibiting the viral spike protein–ACE2 interaction for prophylaxis and treatment in COVID-19 for susceptible pre-Omicron variants[43,44]. However, targeting the viral spike protein with monoclonal antibodies is limited by the diversity and evolution of the viral spike sequences, which renders many of these agents ineffective against new SARS-CoV-2 variants[7]. By contrast, targeting ACE2 is advantageous for several reasons. ACE2 modulation is a host-directed treatment, which does not target the virus. Such mechanisms may present a higher barrier to the emergence of resistance, although this has yet to be empirically demonstrated. Furthermore, because ACE2 is a critical mechanism for cell entry, the approach may be more resilient as variants continue to emerge[45]. Finally, ACE2 is a common receptor for multiple coronaviruses, such as SARS-CoV and HCoV-NL63. Confirmation of the efficacy of this strategy may therefore provide a quickly deployable intervention in the event of future coronavirus outbreaks. Taken collectively with our own observations on the effects of ACE2 modulation for SARS-CoV-2 infection, these points indicate that ACE2 modulators warrant consideration as priority candidates for clinical evaluation in COVID-19 trials[1].

Our finding that the suppression of FXR signalling through UDCA or ZGG reduces ACE2 expression and limits SARS-CoV-2 infection identifies a potential clinical application for FXR inhibitors, but also raises some points

for consideration. First, FXR activation decreases inflammation by modulating NF-κB in several organs[46], including the lungs[21], liver[17] and intestine[17]. Conversely, UDCA has been shown to reduce inflammation in multiple tissues, including the lung, in an FXR-independent manner[27]. Given the complex interplay among FXR, UDCA and inflammation[47], the balance of benefits for FXR activation in terms of SARS-CoV-2 infection and inflammation should be carefully considered. It is possible that FXR suppressors beyond UDCA, which lack anti-inflammatory effects, would be better suited for prophylaxis or early intervention and not indicated for severe disease with ongoing tissue inflammation[48]. Second, our study suggests that FXR activators that are used in clinical practice, such as obeticholic acid, may increase the risk of developing COVID-19 by upregulating ACE2 in healthy individuals. Paradoxically, in patients with liver disorders, obeticholic acid may prevent COVID-19 by reducing disease severity and ameliorating cholestasis, resulting in a net reduction in FXR activity. Therefore, further studies are needed to clarify these points.

Our results identify UDCA as a particularly advantageous modulator of the levels of ACE2 for use in COVID-19. We have shown that UDCA reduces SARS-CoV-2 infection in vitro, in vivo and ex vivo, and that it reduces the expression of ACE2 in the nasal epithelium of healthy volunteers. Although our animal data do not exclude the possibility that UDCA delays the transmission of SARS-CoV-2 beyond the duration of our experiments, our patient data show that this does not change the net effect of reducing disease severity, thus making UDCA a particularly attractive candidate for investigation as pharmacological prophylaxis against SARS-CoV-2 infection. Compared to other agents, such as vaccines and monoclonals, UDCA is easy to administer orally, easily stored, affordable and accessible to health systems worldwide for large-scale production, as it is off patent. In addition, UDCA is well-tolerated and has limited drug–drug interactions and a favourable safety profile, enabling it to be administered for long periods of time. Of note, UDCA is already administered over long periods for different clinical indications to vulnerable groups who would benefit from chemoprophylaxis, such as individuals who have received bone marrow or liver transplants (for the prevention of veno-occlusive disease[49] and for the treatment of cholangiopathy[26], respectively). It has excellent tolerability and minimal side effects in these groups of patients[26,49], which shows the

feasibility of using UDCA as pharmacological prophylaxis against COVID-19 in vulnerable groups. Nevertheless, our study is not a clinical trial and therefore we cannot exclude the potential for confounding and selection biases. Consequently, it will be imperative to validate these results in prospective double-blinded clinical trials and to fully assess the effect of this drug on ACE2 levels and susceptibility to SARS-CoV-2 infection. For the avoidance of doubt, the authors do not support the use of UDCA for COVID-19 until appropriate policy informed by robust clinical evidence is available. The authors also do not condone the use of UDCA as a substitute for highly effective vaccinations in patients for whom they are indicated.

Finally, we have shown that UDCA an reduce ACE2 levels and SARS-CoV-2 infection in machine-perfused organs. This is one of the first studies to test the effect of a drug in a whole human organ perfused ex situ. This finding could prove important for organ transplantation, especially given concerns about peri-operative viral transmission[50]. Furthermore, although more data are required to definitively establish this approach, our work sets the stage for future studies using machine-perfused organs for pharmacological studies.

In conclusion, these results validate CDCA-treated cholangiocyte organoids as a platform for disease modelling and drug testing against SARS-CoV-2 infection, identify FXR as a therapeutic target in the management of COVID-19 and open up new avenues for the modulation of ACE2 through FXR for the prevention of infection with SARS-CoV-2 as well as other viruses that use ACE2 for cell entry.

Methods

Ethical approval

All human samples were obtained from patients, deceased transplant organ donors or liver explants with informed consent for use in research and ethical approval (Research Ethics Committee (REC) 09/H0305/68, 14/NW/1146, 15/EE/0152, 15/WA/0131 and 18/EE/0269; and Papworth Hospital Research Tissue Bank project number T02233). The mouse study was approved by the Animal Ethics Committee of the Medical University of

Vienna and the Federal Ministry of Science, Research and Economy (BMWFW-66.009/0008-WF/3b/2015) and was performed according to the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines. The hamster animal study was approved by the University of Liverpool Animal Welfare and Ethical Review Board and performed under UK Home Office licences (PP9284915 and PP4715265) and it was completed at the University of Liverpool and conducted in accordance with the UK Home Office Animals Scientific Procedures Act (ASPA, 1986). Human lungs and livers retrieved for transplantation but subsequently declined were used for ESNP experiments (National Research Ethics Committee (NREC) North East—Newcastle and North Tyneside 1 16/NE/0230, lung; NREC East of England—Cambridge East 14/EE/0137, liver). The study involving volunteers from the University Medical Centre Hamburg-Eppendorf was performed with informed consent and ethical approval (Ethik-Kommission der Ärztekammer Hamburg; ref. no. 2021-300121-WF). The COVID-Hep.net and SECURE-Liver registries data were deemed not to constitute human research by Clinical Trials and Research Governance at the University of Oxford (https://covid-hep.net/img/CTRG_COVID-Hep_20200402.pdf) and by the Institutional Review Board of University of North Carolina (https://covidcirrhosis.web.unc.edu/faq/), respectively. The study involving patients from the VOCAL cohort was performed with informed consent and ethical approval from the Miami VA Institutional Review Board (unique study approval ID 1477437-22).

10X single-cell RNA sequencing, data analysis and availability

We used our previously published single-cell RNA sequencing (scRNA-seq) dataset including primary cholangiocytes, cholangiocyte organoids (COs) originating from different regions of the biliary tree (intrahepatic ducts, common bile duct and gall bladder) and the same organoids after bile treatment. Tissue dissociation, cell isolation, 10X single-cell library preparation and 10X data processing, normalization and analysis was performed as previously described[14]. The 10X raw data (fastq files) have been deposited in the repository ArrayExpress with the accession number E-MTAB-8495. scRNA-seq data were analysed using Anaconda-Navigator v.1.9.12, Jupyter Notebook v.6.0.3 and Rstudio v.1.1.463.

Human tissue collection and processing

Human primary tissue was obtained from biopsies, deceased organ donors or liver explants after obtaining informed consent. Depending on the application, primary fresh tissue was embedded in optimal cutting temperature (OCT) compound and stored at −80 °C; or fixed in 10% formalin, dehydrated and embedded in paraffin. Sections from embedded tissue were cut at a thickness of 5–10 μm using a cryostat or a microtome and mounted on microscopy slides for further analysis.

Bile sample collection and processing

Human bile was collected during endoscopic retrograde cholangio-pancreatography (ERCP) or intraoperatively with informed consent from the patient. For viral RNA quantification, samples were immediately lysed using an equal volume of RNA lysis buffer (Sigma) and stored at −20 °C.

Cell culture

Primary cholangiocytes were isolated and COs were derived and cultured using our established methodology[14,16,51]. COs obtained from intrahepatic duct (IHD), common bile duct (CBD) and gall bladder (GB) tissue were used in this study. Cholangiocytes derived from any of the different regions of the biliary tree (IHD, CBD or GB) acquired a common gall bladder identity when treated with CDCA, as previously reported[14]. The experiments described were performed with COs derived from all the three regions of the biliary tree (IHD, CBD and GB) and provided congruent results. For consistency, the results shown correspond only to COs derived from the gall bladder (gall bladder cholangiocyte organoids—GCOs).

Human primary intestinal organoids, derived from terminal ileum biopsies, were provided by S.J.A.B.'s group. The organoids were derived following a modification of previously described protocols[52], embedded in Matrigel and cultured in Intesticult (StemCell Technologies) supplemented with penicillin–streptomycin and Rho kinase inhibitor (Stratech Scientific).

Human primary airway organoids were provided by J.-H. L.'s laboratory. The organoids were derived and cultured as previously described[24].

Vero E6 cells (ATCC CRL—1586), HEK293 cells (ATCC CRL—1573) and HEK293T cells (ATCC CRL—3216) were grown on tissue culture plates or T25 flasks in 10% FBS DMEM supplemented with l-glutamine and penicillin–streptomycin as previously described[53]. All cell cultures were routinely tested for mycoplasma.

Availability of biological materials

Detailed protocols for the derivation of primary organoids have been previously reported[24,51]. Cell lines are available from standard commercial sources (https://www.lgcstandards-atcc.org: Vero E6 cells, ATCC CRL—1586; HEK293 cells, ATCC CRL—1573).

Modulation of FXR activity

CDCA and UDCA were purchased from Sigma-Aldrich (C9377-5G and U-5127-5G), and ZGG was purchased from Santa Cruz (sc-204414) and reconstituted following the manufacturer's instructions. To modulate FXR activity, organoids were incubated with a final concentration of 10 μM CDCA, or 10 μM CDCA in combination with 10 μM of UDCA or ZGG.

FXR knockdown

FXR knockdown was performed in COs using commercially available lentiviral particles carrying shRNA gene silencer sequences against the human FXR (*NR1H4*) transcript (Santa Cruz; sc-38848-V). Commercially available lentiviral particles carrying control (scrambled) shRNA sequences (Santa Cruz; sc-108080) were used as control. Successfully transduced COs were selected with puromycin 24 h after viral transduction. Quantification of FXR, ACE2 and SHP expression and SARS-CoV-2 infection was performed 10 days after FXR knockdown.

ChIP

Approximately 6×10^6 cells were used for each ChIP, and cells were incubated with fresh medium with 100 µM of CDCA, UDCA or ZGG 2 h before collection. ChIP was performed using the True Micro ChiP kit (Diagenode C01010130) according to the manufacturer's instructions. In brief, following pre-clearing, the lysate was incubated overnight with the FXR antibody (Santa Cruz sc-25309 X) (Supplementary Table 1) or non-immune IgG. ChIP was completed and immunoprecipitated DNA was purified using MicroChip DiaPure columns (Diagenode C03040001). Samples were analysed by qPCR using the $\Delta\Delta C_t$ approach as previously described[51] (see Supplementary Table 3 for primer sequences). Primers flanking the FXRE on the well-known FXR target gene *OSTα* (also known as *SLC51A*; ref. [54]) were used as a positive control, whereas primers flanking a site distant from the FXRE on the *ACE2* promoter were used as a negative control. The results were normalized to the enrichment observed with non-immune IgG ChIP controls.

Luciferase reporter

Two different fragments containing the FXRE IR-1 in the *ACE2* gene and in the *SHP* gene (also known as *NR0B2*) were amplified using human genomic DNA as a template and inserted onto a pGL3-promoter luciferase vector. The ACE2 and SHP IR-1 mutants were generated using a site-directed mutagenesis approach (New England BioLabs E0554S). Sequences of primers used are reported in Supplementary Table 4. These gene reporter constructs were co-transfected with a commercially available FXR expression plasmid (OriGene, SC329876) into HEK293 cells using TransIT-293 Transfection Reagent (MirusBio). Twenty-four hours after transfection, cells were treated with 50 µM of CDCA, UDCA and ZGG in fresh medium for 8 h. Luciferase activity was determined with the GLO-Luciferase Reporter Assay System (Promega, Madison, ONE-Glo Luciferase Assay System) and values were normalized to the empty pGL3 vector.

Immunofluorescence, RNA extraction and qPCR

Immunofluorescence, RNA extraction and qPCR were performed as previously described[14,15,16]. A complete list of the primary and secondary

antibodies used is provided in Supplementary Table 1. A complete list of the primers used is provided in Supplementary Table 2.

All qPCR data were obtained using a QuantStudio 5 384 Well Block (Thermo Fisher Scientific). All qPCR data are presented as the median, IQR and range (minimum to maximum) of four independent experiments unless otherwise stated. Values are relative to the housekeeping gene hydroxymethylbilane synthase (*HMBS*) or glyceraldehyde-3-phosphate dehydrogenase (*GAPDH)*. Statistical analysis is described in the relevant section.

For comparative immunofluorescence images, the cells or sections being compared were stained simultaneously, using the same primary and secondary antibody master mix. All immunofluorescence images were acquired using a Zeiss LSM 700 or 710 confocal microscope using ZEN 2011 SP7 (Zeiss). The same laser power and exposure settings were used to acquire comparative images. Imagej 2.0.0-rc-69/1.53f software (W. Rasband, http://imagej.nih.gov/ij) was used for image processing. Each immunofluorescence image is representative of at least three different experiments.

Flow cytometry analyses

Flow cytometry in organoids was performed as previously described[51]. In summary, organoids were collected using Cell Recovery Solution (Corning) for 20 min at 4 °C and were then centrifuged at 444*g* for 4 min and dissociated to single cells using StemPro Accutase (Invitrogen). Cells were subsequently fixed using 4% paraformaldehyde (PFA) for 20 min at 4 °C. All flow cytometric analyses were performed on a BD LSR-II flow cytometer (BD Biosciences) using BD FACS Diva 8.0.3 (BD Bioscience) and analysed using FlowJo v.10.4.2. The gating strategy is provided in Supplementary Fig. 1.

Dose–response curves for ACE2

Primary organoids were treated with 0.01 µM–1 mM of CDCA, UDCA or ZGG and ACE2 expression was measured by qPCR. The inhibitory effect of

UDCA and ZGG on FXR activation was assessed on cells treated with 10 μM of CDCA. Data were analysed using the Sigmoidal, 4PL, X is log(concentration) function in GraphPad Prism.

Cytotoxicity and viability

Primary organoids were treated with 0.1 μM–100 μM of CDCA, UDCA or ZGG and the percentage of viable cells was counted using trypan blue and a Countess II cell counter (Thermo Fisher Scientific). Cellular viability in primary organoids treated with 10 μM of CDCA, UDCA or ZGG was measured using the resazurin-based assay PrestoBlue (Invitrogen, A13261) using SoftMax Pro 5.4.4 on a SpectraMax M2 (Molecular Devices).

SARS-CoV-2 isolate

The SARS-CoV-2 virus used in this study is the clinical isolate named SARS-CoV- 2/human/Liverpool/REMRQ0001/2020 (ref. [53]) derived from a patient's nasopharyngeal swab and isolated by L. Turtle, D. Matthews and A. Davidson, and the Delta lineage (B.1.617.2) hCoV-19/England/SHEF-10E8F3B/2021 (GISAID accession number EPI_ISL_1731019) was provided by W. Barclay through the Genotype-to-Phenotype National Virology Consortium (G2P-UK).

SARS-CoV-2 infection

All work with infectious SARS-CoV-2 was performed under containment level 3 (CL-3) conditions either at the Cambridge institute of Therapeutic Immunology and Infectious Diseases (CITIID) or at the Centre for Excellence and Long-acting Therapeutics (CELT). SARS-CoV-2 was gifted to the users of the CITIID CL-3 by I. Goodfellow[55,56] and propagated on Vero E6 cells as previously described[53]. Viral titration was determined using the TCID50 method on Vero E6 cells[53]. For viral infection primary organoids were passaged and incubated with SARS-CoV-2 in suspension at a multiplicity of infection (MOI) of 1 for 2 h. Subsequently, the infected organoids were washed twice with 10 ml of culture medium to remove the viral particles. Washed organoids were plated in 40-μl Matrigel domes, cultured in organoid medium and collected at different time points.

To test whether SARS-CoV-2 produced by infected COs retained its infective capacity, the supernatant from infected COs was collected 24 h after infection and used to infect a fresh batch of SARS-CoV-2 naive organoids.

Fixation of SARS-CoV-2 infected organoids or tissue

Organoids for immunofluorescence were cultured on coverslips and placed at the bottom of the wells of a 24-well plate. The culture medium was aspirated and replaced with 500 µl of 8% PFA for a minimum of 30 min. After fixation, the coverslips were recovered and transferred to a clean plate, and fresh PBS was added. Primary tissue was fixed for a minimum of 4 h with 8% PFA and then transferred to a clean plate with fresh PBS.

Quantification of viral infection

Organoids or primary tissue were infected in 24-well plates as described above. Total RNA samples were prepared by adding 500 µl of lysis buffer (25 mM Tris-HCl + 4 M guanidine thiocyanate with 0.5% β-mercaptoethanol) to each well and transferring the lysate (1 ml) to a 5-ml Eppendorf tube. Tubes were vortexed, and 100% analytical grade ethanol was added to a final concentration of 50%. After 10 min of incubation, 860 µl of lysis buffer (containing MS2 bacteriophage as an internal extraction and amplification control) was added and thoroughly mixed. The RNA was then isolated using an RNA spin column as previously described[57]. Viral replication was quantified using qPCR for the expression of the viral RNA-dependent RNA polymerase (*RdRp*) gene with primers specific for a 222-bp fragment from a conserved region of the gene. *GAPDH* was used as a housekeeping gene and MS2 was used as an internal reference as previously described[57]. Viral load was determined relative to *GAPDH*. The sequences of primers and probes used are provided in Supplementary Table 2.

Transmission electron microscopy

Infected organoids were fixed in 4% paraformaldehyde; 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer overnight at 4 °C, washed and stored in 0.1 M sodium cacodylate buffer before processing. Samples

were post-fixed in 1% aqueous osmium tetroxide (TAAB) and 1.5% potassium ferricyanide overnight at 4 °C, washed thoroughly in dH_2O and en-bloc-stained in 3% aqueous uranyl acetate (Agar Scientific) for 24 h at 4 °C. Samples were dehydrated through an ethanol series, infiltrated with 1:1 propylene oxide:resin (TAAB) and blocks of fresh resin polymerized at 60 °C for 48 h. Ultrathin sections of around 60 nm were cut from blocks using an EM UC7 ultramicrotome (Leica Microsystems) and mounted on copper grids coated with carbon and formvar (Agar Scientific). Grids were post-stained in uranyl acetate and lead citrate, imaged using a HT7800 transmission electron microscope (Hitachi High Technologies) operating at 100 kV and acquired using HT7800 TEM operating software v.01.21 (Hitachi).

HEK293 cells stably expressing ACE2

HEK293T cells stably expressing ACE2 were generated as previously described[28]. In brief, HEK293T cells transduced with ACE2 under the control of the spleen focus-forming virus (SFFV) promoter were sorted for high cell-surface ACE2 expression and single-cell-cloned. After expansion, a clone with stable, homogeneously high expression of ACE2 was selected by fluorescence-activated cell sorting (FACS).

Luciferase reporter for SARS-CoV-2 replication

A luciferase reporter for SARS-CoV-2 protease activity during viral replication was generated as previously described[28] In brief, HEK293T reporter cells stably expressing ACE2, renilla luciferase (Rluc) and SARS-CoV-2 papain-like protease-activatable circularly permuted firefly luciferase (FFluc) were seeded in flat-bottomed 96-well plates. The following morning, cells were treated with the indicated doses of CDCA, UDCA and ZGG, and infected with SARS-CoV-2 at a MOI of 0.01. The SARS-CoV-2 RdRp inhibitor remdesivir and a neutralizing antibody cocktail blocking the interaction between SARS-CoV-2 spike and ACE2 (REGN-COV2) were included as positive controls. After 24 h, cells were lysed in Dual-Glo Luciferase Buffer (Promega, E2920) diluted 1:1 with PBS and 1% NP-40. Lysates were then transferred to opaque 96-well plates, and viral replication quantified as the ratio of FFluc/Rluc activity measured using the Dual-Glo

kit (Promega) according to the manufacturer's instructions. FFluc/Rluc ratios were expressed as a fraction of the maximum, then analysed using the Sigmoidal, 4PL, X is log(concentration) function in GraphPad Prism.

Animal studies

Mice

The experiments were performed in accordance with the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines and approved by the Animal Ethics Committee of the Medical University of Vienna and the Federal Ministry of Science, Research and Economy (BMWFW-66.009/0008-WF/3b/2015). Friend Virus B NIH (FVB/N) mice were bred in house. Mice were housed in a 12 h–12 h dark–light cycle, with a humidity of 45–65% and temperature of 20–24 °C. Chow was obtained from SAFE—Scientific Animal Food & Engineering (product number A04). Age-matched female mice were used. Mice were assigned randomly to treatment and control groups. Mice in the treatment group received chow supplemented with 1% w/w UDCA and 1% w/w cholic acid, whereas mice in the control group received chow supplemented with 1% w/w cholic acid[58]. Cholic acid was used to activate FXR and study the effects of UDCA on FXR activation[19]. The mice were fed ad libitum for seven days. Data were analysed blinded to the identity of the experimental groups.

Hamsters

The experiments were performed in accordance with the UK Home Office Animals Scientific Procedures Act (ASPA, 1986). In addition, all studies were approved by the University of Liverpool Animal Welfare and Ethical Review Board and performed under UK Home Office licences PP9284915 and PP4715265. Golden Syrian hamsters were purchased from Janvier Labs. Hamsters were housed in a 12 h–12 h dark–light cycle, with a humidity of 45–65% and temperature of 20–24 °C. Age-matched male hamsters were used, weighing between 80 g and 100 g. Hamsters were assigned randomly to treatment and control groups. Hamsters in the treatment groups received a daily oral regimen of UDCA (416 mg per kg) by oral gavage, whereas those

in the control group received vehicle only. The hamsters were fed ad libitum and treatment continued for seven days to achieve a similar blood concentration of UDCA to that observed in patients taking UDCA[29] (Extended Data Fig. 9a).

For testing the effects of UDCA against SARS-CoV-2 infection, one hamster was directly inoculated by the intranasal route with 1×10^2 plaque-forming units (PFU) in 100 µl PBS. Each infected hamster was placed on one side of a transmission cage. The cage was divided with an aerated barrier that allowed the infected hamster to be co-housed with previously treated uninfected hamsters housed on the other side, permitting us to study viral infection by aerosol transmission. Daily swabs were collected from all hamsters to monitor the infection by qPCR for the viral N gene. On day 4 after infection, the hamsters were euthanized and lungs and nasal turbinates were collected for quantification of viral infection. The experiment was repeated $n = 3$ times for a total of $n = 9$ UDCA and n = 6 vehicle hamsters. Data were analysed blinded to the identity of the experimental groups.

Quantification of UDCA concentration

UDCA was quantified from hamster plasma using a liquid chromatography–mass spectrometry (LC–MS) assay that was validated using FDA industry guidelines. Quantification was achieved using LC–MS/MS (6500+ QTRAP, SCIEX) operating in negative mode. UDCA was detected using multiple reaction monitoring (MRM) in which the following ions were monitored for quantification: UDCA (*m*/*z* 391>391 and internal standard mefloquine 379.1>320.1). A stock solution of 1 mg ml^{-1} UDCA was prepared in methanol and stored at 4 °C until use. A standard curve was prepared in plasma by serial dilution from 40,000 ng ml^{-1} to 312.5 ng ml^{-1} and an additional blank solution was also used. Chromatographic separation was achieved using a multi-step gradient with a Acquity BEH C18 column (2.1 mm × 100 mm 1.7 µm; Waters) using mobile phases A (100% water, 0.1% formic acid and 5 mM ammonium formate) and B (90% acetonitrile 10% methanol, 0.1% formic acid and 5 mM ammonium formate). Chromatography was conducted over 3.5 min. At the start of each run, mobile phase A was 80% until 0.5 min, when mobile phase B was increased to 47% over 0.5 min. Mobile phase B was then increased over 1 min to 51%.

Mobile phase B was then increased to 100% at 2.5 min, which was held until 3 min. Mobile phase B was reduced to 20% and held until 3.5 min. Samples were extracted from hamster plasma by protein precipitation. In brief, 100 µl of standard, quality control, blank plasma or study sample were treated with 400 µl of acetonitrile. Samples were then vortexed followed by centrifugation at 3,500 rpm for 5 min, and 400 µl of supernatant was transferred to fresh glass vials and evaporated under a steady stream of nitrogen. Samples were reconstituted in 50:50 water methanol and analysed. Inter- and intra- assay variance was assessed by three levels of independent quality controls. The coefficients of variation of accuracy and precision were less than 15% in all assays.

ESNP of human lungs

For the ex situ perfusion of a single pair of human lungs, two bespoke ESNP circuits (Medtronic) were used. In brief, this circuit facilitates pressure-monitored perfusion with normothermic perfusion solution consisting of bovine serum albumin (BSA) (70 g l^{-1}), dextran 40 (5 g l^{-1}), modified Krebs Henseleit buffer (9.2 g l^{-1}), sodium bicarbonate solution (28 ml l^{-1}), calcium chloride (25 ml l^{-1}) and heparin (3,750 units per l). The pair of human lungs used was perfused following the physiological principles previously described[59].

Lung ESNP experimental set-up

The experiment shown in Fig. 3 was performed on a pair of lungs that was declined for clinical lung transplantation owing to the donor's past medical history. The left and the right lungs were divided at the carina and common pulmonary artery bifurcation. For each isolated lung, the pulmonary artery and the pulmonary vein were cannulated and an endobronchial tube was inserted into the main bronchus. The cannulae of each lung were connected to an entirely independent ESNP circuit (control and experimental lung circuits). The endobronchial tube of each lung was connected to an independent mechanical ventilator (Drägerwerk AG & Co. KGaA). Mechanical ventilation was performed using room air with a positive end expiratory pressure of 5 mmHg and a target tidal volume of 130 ml at 5 bpm for the left lung and 140 ml at 6 bpm for the right lung.

Timing and duration of ESNP

Perfusion started simultaneously for both lungs. After 30 min of stable perfusion, UDCA or carrier was administered and this time point was defined as 0 h (0 h). ESNP was performed for 6 h for both lungs after administration of UDCA or carrier. The end of experiment time point is defined as 6 h (6 h).

Experimental time points and sample collection

Baseline tissue and perfusate samples were collected simultaneously from both lungs before UDCA or carrier administration. We defined these samples as 0-h samples. Corresponding samples were collected simultaneously from each lung at the end of the experiment at 6 h. We defined those as 6-h samples.

The following samples were collected from each lung per time point: $n = 4$ independent lung biopsies, $n = 4$ surgically excised independent samples from the pulmonary artery, $n = 4$ surgically excised independent samples from the main bronchus and $n = 4$ independent samples from the perfusate in each circuit. The independent tissue samples refer to different locations of the organ. Biopsies were surgically excised and the lung parenchymal samples from 0 h were taken from peripheral areas separated from the remaining lung by staples (Medtronic) to seal the defect. The 6-h samples were taken using the same approach. Please also refer to the main-text section ‘FXR regulates infection in human organs’.

ESNP of donor livers

The OrganOx metra normothermic liver perfusion device was used for ex situ perfusion of human livers as previously described[14,30,32]. The machine, which is clinically used for preservation of livers for transplantation, enables prolonged automated organ preservation by perfusing it with ABO-blood group-matched normothermic oxygenated blood. The perfusion device incorporates online blood gas measurement, as well as software-controlled algorithms to maintain pH, PO2 and PCO2 (within physiological limits),

temperature, mean arterial pressure and inferior vena cava pressure within physiological normal limits.

Experimental set-up

Two donor livers not used for transplantation were maintained with ESNP. In brief, the hepatic artery, portal vein, inferior vena cava and bile duct were cannulated and connected to the device, and perfusion commenced. One liver was randomly chosen to receive a solution of UDCA dissolved in 0.9% NaCl (experimental liver), whereas the other liver (control) was chosen to receive the same volume of carrier (0.9% NaCl). UDCA was resuspended in 0.9% (w/v) NaCl solution and injected in the blood circuit to achieve a final concentration of 2,000 ng ml^{-1}, which is the steady-state concentration of UDCA detected in serum after multiple doses of UDCA[29].

Timing and duration of ESNP

The time of UDCA or carrier administration after the start of ESNP was defined as 0 h (0 h). ESNP was performed for 12 h for both livers after administration of UDCA or carrier. The end of experiment time point is defined as 12 h (12 h).

Experimental time points and sample collection

Baseline tissue and perfusate samples were collected from each liver before UDCA or carrier administration. We defined these samples as 0 h samples. Corresponding samples were collected from each liver at the end of the experiment at 12 h. We defined those as 12 h samples.

The following samples were collected from each liver per time point: $n = 4$ independent surgically excised gall bladder samples, $n = 4$ independent liver biopsies and $n = 4$ independent samples from the circulating perfusate. Independent gall bladder samples and liver biopsies were obtained from different locations of the organ. Gall bladder samples were surgically excised and the gall bladder wall was sutured to close the defect.

Ex vivo infection of human tissue with SARS-CoV-2

The infection of human tissue maintained ex vivo with SARS-CoV-2 occurred in a CL-3 facility after the ESNP experiment. Four independent samples from the lung parenchyma, bronchi and the vasculature (pulmonary artery) of each lung were collected per time point; that is, at the start of the ESNP experiment (0 h) and at the end (6 h) of the lung perfusion. Four independent samples from each ESNP liver gall bladder were collected per time point; that is, at the start of the ESNP experiment (0 h) and at the end (12 h) of the liver perfusion. Sample collection was performed as described in 'ESNP of donor livers'.

The freshly obtained lung, bronchial, vascular and gall bladder tissues were processed into small rectangular pieces of 0.5 × 0.5 cm and were rinsed with University of Wisconsin solution (lung, bronchial and vascular tissues) or William's E medium with supplements as previously described[51] (gall bladder tissue). Washed specimens were placed in wells of a 24-well plate (one specimen per well) and infected with SARS-CoV-2. An inoculum of 1.2×10^5 PFU per ml at 500 µl per well was used. After two hours, the inoculum was removed, and the specimens were washed three times with PBS. The infected human tissue was then cultured in 500 µl of advanced DMEM with supplements[60] (lung, bronchial and vascular tissues) or

William’s E medium with supplements (gall bladder tissue). Supernatant and tissue were collected for qPCR and immunofluorescence at 2 h and 24 h after infection.

Blood sample collection and processing

Blood samples were collected from patients as part of the UK-PBC Nested Cohort study after obtaining informed consent, anonymized and analysed by a blinded researcher. To obtain serum from full blood, the samples were spun at 4 °C at 1,000*g* for 10 min to allow for serum separation and serum was collected as the supernatant. All blood samples were collected after fasting.

ACE2 enzymatic activity

ACE2 enzymatic activity was performed on serum samples and tissue lysates using the ACE2 activity fluorometric kit (Abcam ab273297) following the manufacturer’s instructions using SoftMax Pro 5.4.4 on a SpectraMax M2 (Molecular Devices).

ACE2 measurement in nasal epithelial cells of volunteers

Recruitment

Following approval by the local ethics committee (Ethik-Kommission der Ärztekammer Hamburg; ref. no. 2021-300121-WF), the study was advertised in the University Medical Centre Hamburg-Eppendorf amongst clinicians who regularly prescribe UDCA, and are thus familiar with the drug and its possible side effects. Eight clinicians who volunteered to participate in the study were recruited after informed consent. The characteristics of the volunteers are provided in Supplementary Table 5.

Study design and exclusion criteria

UDCA was self-administered at the clinically approved dose of 15 mg per kg per day in a single morning dose for 5 days. Nasal epithelium samples

for ACE2 measurement were collected each morning before UDCA administration using the Citoswab nasopharyngeal swab collection kit (Corona Smear). Day 0 samples were collected before the first UDCA dose. Daily morning samples were collected during UDCA treatment (days 1–5). After drug washout, repeat nasopharyngeal swabs were collected between days 22 and 28. Volunteers providing samples with no detectable RNA were excluded from the study.

ACE2 measurement

RNA was extracted from the Citoswab nasopharyngeal swab collection kit (Corona Smear) using the RNeasy micro kit (Qiagen) according to the manufacturer's instructions. ACE2 mRNA levels were measured using qPCR (see 'Immunofluorescence, RNA extraction and qPCR'). The results are shown as fold change over the housekeeping gene *GAPDH*.

Serum proteome analysis in patients with PBC

Serum proteomic analysis was performed in the UK-PBC patient cohort as described previously[33]. Blood samples were obtained with informed consent and appropriate ethical approval (UK-PBC tissue bioresource, 14/NW/1146). Serum samples were assayed using the Olink proteomics platform (https://olink.com/).

Patient data

Data for patients with chronic liver disease were collected as described elsewhere[35,61]. In brief, collated results from two open online reporting registries (COVID-Hep.net and SECURE-Liver) were examined. Reports were asked to record cases of laboratory-confirmed COVID-19 in patients with chronic liver disease at the end of the disease course. Anonymous clinical and demographic data were collected and filtered to remove duplicate entries and entries with incomplete records, and to remove data on individuals with prior liver transplantation, who were not over 18 years of age, who were over 90 years of age or who did not have a laboratory-confirmed infection.

Data for patients with liver transplants were collected as described elsewhere[36]. In brief, data from patients with a liver transplant who were alive on 1 March 2020 were examined. Participants who had no COVID-19 infection, were unvaccinated or developed COVID-19 within 30 days of their first UDCA prescription were excluded. The resulting study sample included $n = 24$ patients on UDCA and 95 who were not on UDCA, and this sample was used for the analysis.

Statistical analyses

Statistical analyses were performed using Microsoft Excel v.16.19, Rstudio (v.1.1.463), GraphPad Prism 9 or Stata 15.1 (StataCorp). The normal distribution of our values was evaluated using the Shapiro-Wilk test where appropriate. For comparison between two groups, a two-tailed Student's *t*-test or the non-parametric Mann–Whitney test was used depending on the normality of our distribution. To compare matched samples from the same individual, a two-tailed paired Student's *t*-test was used (the normal distribution of our data was confirmed using the Shapiro-Wilk test). Variance between samples was tested using the Brown-Forsythe test. For comparing multiple groups to a reference group, a one-way ANOVA followed by Dunnett's test was used between groups with equal variance. Immunofluorescence images are representative of four independent experiments. Data are represented in box plots and elements are defined as follows: centre line, median; box, IQR; whiskers, range; error bars, s.d.

Serum proteomic analysis was performed using Rstudio (v.1.1.463). The correlation between ACE2 levels and UDCA administration was interrogated using multiple linear regression analysis with the lm function in R. ACE2 expression data was defined as the independent variable, and UDCA administration, sex, age, BMI, stage of chronic liver disease according to the Child–Turcotte–Pugh class and alkaline phosphatase were defined as dependent variables.

COVID-Hep and SECURE-Liver data (exploratory cohort): for propensity-score-matched analyses, 1:5 matched samples (using the nearest-neighbour approach) were constructed with hospitalization, physician-reported requirement for intensive care, intensive care admission, mechanical

ventilation and death as the outcome variables. The covariables used were age, sex and categorical stage of chronic liver disease according to the Child–Turcotte–Pugh class, diabetes, chronic pulmonary disease (COPD or asthma), immunosuppressive therapy, increased BMI (BMI > 25), and the presence of non-alcoholic fatty liver disease, owing to its association with increased COVID-19 risk. Notably, ARLD is also associated with increased COVID-19 risk; however, patients could not be propensity-score-matched between the two cohorts, as there were no patients with ARLD receiving UDCA. Consequently, patients with ARLD were excluded.

VOCAL data (validation cohort): for propensity-score-matched analyses, 1:3 matched samples (using a Greedy matching algorithm) were constructed with moderate, severe or critical COVID-19 and severe or critical COVID-19 according to the NIH severity scale as the outcome variables. The covariables used were age, sex, ethnicity, location within the United States, diabetes, BMI, COPD, type of immunosuppressive therapy (calcineurin inhibitor with or without anti-metabolite therapy) and dominant SARS-CoV-2 variant.

Propensity-score matching was performed using the teffects function in Stata. The average treatment effect on the treated (ATET) was calculated with robust Abadie-Imbens standard errors and derived 95% CIs are presented.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

scRNA-seq data are available on ArrayExpress under accession number E-MTAB-8495. Source data are provided with this paper.

References

1. Gaziano, L. et al. Actionable druggable genome-wide Mendelian randomization identifies repurposing opportunities for COVID-19. *Nat. Med.* **27**, 668–676 (2021).

2. World Health Organization. *WHO Guidelines: Drugs to Prevent COVID-19* https://www.who.int/publications/i/item/WHO-2019-nCoV-prophylaxes-2021-1 (WHO, 2021).

3. Bartoszko, J. J. et al. Prophylaxis against covid-19: living systematic review and network meta-analysis. *Br. Med. J.* **373**, n949 (2021).

4. World Health Organization. Therapeutics and COVID-19: living guideline (version 9.3, 3 March 2022) (WHO, 2022).

5. The RECOVERY Collaborative Group. Dexamethasone in hospitalized patients with Covid-19. *N. Engl. J. Med.* **384**, 693–704 (2021).

6. Beigel, J. H. et al. Remdesivir for the treatment of Covid-19—final report. *N. Engl. J. Med.* **383**, 1813–1826 (2020).

7. Cao, Y. et al. Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature* **602**, 657–663 (2022).

8. Collier, D. A. et al. Sensitivity of SARS-CoV-2 B.1.1.7 to mRNA vaccine-elicited antibodies. *Nature* **593**, 136–141 (2021).

9. Dyer, O. Covid-19: countries are learning what others paid for vaccines. *Br. Med. J.* **372**, n281 (2021).

10. Callaway, E. The unequal scramble for coronavirus vaccines—by the numbers. *Nature* **584**, 506–507 (2020).

11. Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).

12. Sungnak, W. et al. SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat. Med.* **26**, 681–687 (2020).

13. Evangelou, K. et al. Pulmonary infection by SARS-CoV-2 induces senescence accompanied by an inflammatory phenotype in severe COVID-19: possible implications for viral mutagenesis. *Eur. Respir. J.* **60**, 2102951 (2022).

14. Sampaziotis, F. et al. Cholangiocyte organoids can repair bile ducts after transplantation in the human liver. *Science* **371**, 839–846 (2021).

15. Tysoe, O. C. et al. Isolation and propagation of primary human cholangiocyte organoids for the generation of bioengineered biliary tissue. *Nat. Protoc.* **14**, 1884–1925 (2019).

16. Sampaziotis, F. et al. Reconstruction of the mouse extrahepatic biliary tree using primary human extrahepatic cholangiocyte organoids. *Nat. Med.* **23**, 954–963 (2017).

17. Sun, L., Cai, J. & Gonzalez, F. J. The role of farnesoid X receptor in metabolic diseases, and gastrointestinal and liver cancer. *Nat. Rev. Gastroenterol. Hepatol.* **18**, 335–347 (2021).

18. Urizar, N. L. et al. A natural product that lowers cholesterol as an antagonist ligand for FXR. *Science* **296**, 1703–1706 (2002).

19. Sun, L. et al. Gut microbiota and intestinal FXR mediate the clinical benefits of metformin. *Nat. Med.* **24**, 1919–1929 (2018).

20. Chen, B. et al. Bile acids induce activation of alveolar epithelial cells and lung fibroblasts through farnesoid X receptor-dependent and independent pathways. *Respirology* **21**, 1075–1080 (2016).

21. Comeglio, P. et al. Anti-fibrotic effects of chronic treatment with the selective FXR agonist obeticholic acid in the bleomycin-induced rat model of pulmonary fibrosis. *J. Steroid Biochem. Mol. Biol.* **168**, 26–37 (2017).

22. Fickert, P. & Wagner, M. Biliary bile acids in hepatobiliary injury—what is the link? *J. Hepatol.* **67**, 619–631 (2017).

23. Gupta, A. et al. Extrapulmonary manifestations of COVID-19. *Nat. Med.* **26**, 1017–1032 (2020).

24. Youk, J. et al. Three-dimensional human alveolar stem cell culture models reveal infection response to SARS-CoV-2. *Cell Stem Cell* **27**, 905–919 (2020).

25. Lamers, M. M. et al. SARS-CoV-2 productively infects human gut enterocytes. *Science* **3**, 50–54 (2020).

26. EASL Clinical Practice Guidelines.The diagnosis and management of patients with primary biliary cholangitis. *J. Hepatol.* **67**, 145–172 (2017).

27. Sharma, R. et al. Ursodeoxycholic acid amides as novel glucocorticoid receptor modulators. *J. Med. Chem.* **54**, 122–130 (2011).

28. Gerber, P. P. et al. A protease-activatable luminescent biosensor and reporter cell line for authentic SARS-CoV-2 infection. *PLoS Pathog.* **18**, e1010265 (2022).

29. Lee, S. et al. Pharmacokinetics of ursodeoxycholic acid in elderly volunteers compared with younger adults in a Korean population. *J. Clin. Pharmacol.* **59**, 1085–1092 (2019).

30. Nasralla, D. et al. A randomized trial of normothermic preservation in liver transplantation. *Nature* **557**, 50–56 (2018).

31. Andreasson, A. S. I., Dark, J. H. & Fisher, A. J. Ex vivo lung perfusion in clinical lung transplantation—state of the art. *Eur. J. Cardiothorac. Surg.* **46**, 779–788 (2014).

32. Watson, C. J. E. et al. Observations on the ex situ perfusion of livers for transplantation. *Am. J. Transplant.* **18**, 2005–2020 (2018).

33. Barron-Millar, B. et al. The serum proteome and ursodeoxycholic acid response in primary biliary cholangitis. *Hepatology* **74**, 3269–3283 (2021).

34. Webb, G. J. et al. Outcomes following SARS-CoV-2 infection in liver transplant recipients: an international registry study. *Lancet Gastroenterol. Hepatol.* **5**, 1008–1016 (2020).

35. Marjot, T. et al. Outcomes following SARS-CoV-2 infection in patients with chronic liver disease: an international registry study. *J. Hepatol.* **74**, 567–577 (2021).

36. John, B. V. et al. Effectiveness of COVID-19 viral vector Ad.26.COV2.S vaccine and comparison with mRNA vaccines in cirrhosis. *Clin. Gastroenterol. Hepatol.* **20**, 2405–2408 (2022).

37. John, B. V. et al. Ursodeoxycholic acid use and outcomes of Coronavirus 2019 in patients with liver disease. *Hepatology* **76**, S557–S558 (2022).

38. Patel, K. et al. Cilofexor, a nonsteroidal FXR agonist, in patients with noncirrhotic NASH: a phase 2 randomized controlled trial. *Hepatology* **72**, 58–71 (2020).

39. Jiang, C. et al. Intestinal farnesoid X receptor signaling promotes nonalcoholic fatty liver disease. *J. Clin. Invest.* **125**, 386–402 (2015).

40. Zhao, K. et al. Activation of FXR protects against renal fibrosis via suppressing Smad3 expression. *Sci. Rep.* **6**, 37234 (2016).

41. Gadaleta, R. M. et al. Farnesoid X receptor activation inhibits inflammation and preserves the intestinal barrier in inflammatory bowel disease. *Gut* **60**, 463–472 (2011).

42. Caron, S., Cariou, B. & Staels, B. FXR: more than a bile acid receptor? *Endocrinology* **147**, 4022–4024 (2006).

43. Weinreich, D. M. et al. REGN-COV2, a neutralizing antibody cocktail, in outpatients with covid-19. *N. Engl. J. Med.* **384**, 238–251 (2021).

44. Levin, M. J. et al. Intramuscular AZD7442 (tixagevimab–cilgavimab) for prevention of Covid-19. *N. Engl. J. Med.* **386**, 2188–2200 (2022).

45. Kemp, S. A. et al. SARS-CoV-2 evolution during treatment of chronic infection. *Nature* **592**, 277–282 (2021).

46. Fiorucci, S., Biagioli, M., Zampella, A. & Distrutti, E. Bile acids activated receptors regulate innate immunity. *Front. Immunol.* **13**, 1853 (2018).

47. Fuchs, C. D. & Trauner, M. Role of bile acids and their receptors in gastrointestinal and hepatic pathophysiology. *Nat. Rev. Gastroenterol. Hepatol.* **19**, 432–450 (2022).

48. Vabret, N. et al. Immunology of COVID-19: current state of the science. *Immunity* **52**, 910–941 (2020).

49. Mohty, M. et al. Prophylactic, preemptive, and curative treatment for sinusoidal obstruction syndrome/veno-occlusive disease in adult patients: a position statement from an international expert group. *Bone Marrow Transplant.* **55**, 485–495 (2020).

50. Qin, J. et al. Perioperative presentation of COVID-19 disease in a liver transplant recipient. *Hepatology* **72**, 1491–1493 (2020).

51. Tysoe, O. C. et al. Isolation and propagation of primary human cholangiocyte organoids for the generation of bioengineered biliary tissue. *Nat. Protoc.* **14**, 1884–1925 (2019).

52. Bjerknes, M. & Cheng, H. Methods for the isolation of intact epithelium from the mouse intestine. *Anat. Rec.* **199**, 565–574 (1981).

53. Guido, P. et al. Furin cleavage of SARS-CoV-2 Spike promotes but is not essential for infection and cell–cell fusion. *PLoS Pathog.* **17**, e1009246 (2021).

54. Balasubramaniyan, N., Luo, Y., Sun, A. Q. & Suchy, F. J. SUMOylation of the farnesoid X receptor (FXR) regulates the expression of FXR target genes. *J. Biol. Chem.* **288**, 13850–13862 (2013).

55. Patterson, E. I. et al. Methods of inactivation of SARS-CoV-2 for downstream biological assays. *J. Infect. Dis.* **222**, 1462–1467 (2020).

56. Cantuti-Castelvetri, L. et al. Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. *Science* **370**, 856–860 (2020).

57. Sridhar, S. et al. A blueprint for the implementation of a validated approach for the detection of SARS-Cov2 in clinical samples in academic facilities. *Wellcome Open Res.* **5**, 110 (2020).

58. Fickert, P. et al. Effects of ursodeoxycholic and cholic acid feeding on hepatocellular transporter expression in mouse liver. *Gastroenterology* **121**, 170–183 (2001).

59. Morrison, M. I. et al. Use of phosphodiesterase inhibition during ex-vivo lung perfusion of donor lungs unsuitable for transplantation. *J. Heart Lung Transplant.* **38**, S321 (2019).

60. Sachs, N. et al. Long-term expanding human airway organoids for disease modeling. *EMBO J.* **38**, e100300 (2019).

61. Marjot, T. et al. SARS-CoV-2 infection in patients with autoimmune hepatitis. *J. Hepatol.* **74**, 1335–1343 (2021).

Acknowledgements

We thank the European Association for the Study of the Liver (EASL) and the American Association for the Study of Liver Disease (AASLD) for supporting the COVID-Hep and SECURE-Liver registries; S. Marciniak and P. J. Lehner for comments and feedback on the manuscript; I. Goodfellow for providing the viral isolate; M. Wills and S. Clare for all their work ensuring a safe CL-3 working environment; C. Cormie for general lab support; the NIHR Cambridge BRC Cell Phenotyping Hub for their help with flow cytometry and processing of samples; the building staff of the Jeffrey Cheah Biomedical Centre for maintaining the institute open and safe during the period of lockdown; K. Füssel for coordinating the volunteer study and sample collection at the University Medical Centre

Hamburg-Eppendorf; J. Hails, K.-I. Nikitopoulou and A. Ford for collecting blood samples; M. Colzani for advising on flow cytometry; A. Wiblin for advising on antibodies; and the Cambridge Biorepository for Translational Medicine for the provision of human tissue used in the study. T.B. was supported by an EASL Juan Rodès PhD fellowship. F.S. was supported by a UKRI Future Leaders fellowship, the Evelyn trust, an NIHR Clinical Lectureship, the Academy of Medical Sciences Starter Grant for Clinical Lecturers, the Addenbrooke's Charitable Trust and the Rosetrees Trust. In addition, the F.S. laboratory is supported by the Cambridge University Hospitals National Institute for Health Research Biomedical Research Centre and the core support grant from the Wellcome Trust and Medical Research Council (MRC) of the Wellcome–Medical Research Council Cambridge Stem Cell Institute. The L.V. laboratory is funded by the ERC advanced grant New-Chol, the Cambridge University Hospitals National Institute for Health Research Biomedical Research Centre and the core support grant from the Wellcome Trust and MRC of the Wellcome–Medical Research Council Cambridge Stem Cell Institute. M.M., S.F. and G.D. are funded by the NIHR Cambridge Biomedical Research Centre and NIHR AMR Research Capital Funding Scheme (NIHR200640). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. V.L.M. was funded by an MRC Clinical Research Training Fellowship. G.F.M. was funded by a post-doctoral fellowship from the National Institute for Health Research (NIHR) Rare Diseases–Translational Research Collaboration (RD-TRC) and by an MRC Clinical Academic Research Partnership (CARP) award. The UK-PBC Nested Cohort study was funded by an MRC Stratified Medicine award (MR/L001489/1). C.J.R.I. was supported by the Medical Research Council (MC_UU_12014). T.M. is funded by a Wellcome Trust Clinical Research Training Fellowship (102176/B/13/Z). The A.P.D. laboratory was supported by BHF TG/18/4/33770, Wellcome Trust 203814/Z/16/A and Addenbrooke's Charitable Trust. The COVID-Hep.net registry was supported by the European Association for the Study of the Liver (EASL) and the SECURE-Liver registry was supported by the American Association for the Study of Liver Disease (AASLD). The lung perfusion experiment was supported by the National Institute for Health Research Blood and Transplant Research Unit (NIHR BTRU) in Organ Donation and Transplantation at Newcastle University and the University of Cambridge

in partnership with NHS Blood and Transplant (NHSBT). The views expressed are those of the author(s) and not necessarily those of the NIHR, the Department of Health and Social Care or NHSBT. G.B. is funded by the European Reference Network for Hepatological Diseases (ERN RARE LIVER). A.O. acknowledges funding for preclinical research on treatment and prevention of COVID-19 from Unitaid (2020-38-LONGEVITY), the Engineering and Physical Sciences Research Council (EPSRC; EP/R024804/1), the Wellcome Trust (222489/Z/21/Z) and UK Research and Innovation (UKRI; BB/W010801/1). N.J.M. acknowledges funding from the MRC (CSF ref. MR/P008801/1 to N.J.M.), NHSBT (grant ref. WPA15-02 to N.J.M.) and Addenbrooke's Charitable Trust (grant ref. to 900239 N.J.M.). This research was funded in whole, or in part, by the Wellcome Trust (203151/Z/16/Z, 203151/A/16/Z) and the UKRI Medical Research Council (MC_PC_17230). For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Author information

Author notes

1. These authors jointly supervised this work: Ludovic Vallier, Fotios Sampaziotis

Authors and Affiliations

1. Wellcome–MRC Cambridge Stem Cell Institute, Cambridge, UK

   Teresa Brevini, Stephanie Brown, Scott Dillon, Sagar S. Varankar, Mahnaz Darvish-Damavandi, James A. Heslop, Davide Rossetti, Olivia C. Tysoe, Vasileios Galanakis, Marta Vila-Gonzalez, Johannes Bargehr, Sanjay Sinha, Darran Clements, Peter Humphreys, Simon J. A. Buczacki, Joo-Hyeon Lee, Ludovic Vallier & Fotios Sampaziotis

2. Cambridge Institute of Therapeutic Immunology & Infectious Disease (CITIID), Department of Medicine, University of Cambridge,

Cambridge, UK

Mailis Maes, Pehuén Pereyra-Gerber, Sally Forrest, Petra Mlcochova, Thomas W. M. Crozier, Ravindra K. Gupta, Stephen Baker, Nicholas J. Matheson & Gordon Dougan

3. Cambridge Liver Unit, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

Gwilym J. Webb, William T. H. Gelson, George F. Mells & Fotios Sampaziotis

4. Division of Gastroenterology and Hepatology, University of Miami and Miami VA Health System, Miami, FL, USA

Binu V. John, Petra Mlcochova & Raphaella D. Ferreira

5. Hans Popper Laboratory of Molecular Hepatology, Division of Gastroenterology and Hepatology, Department of Internal Medicine III, Medical University of Vienna, Vienna, Austria

Claudia D. Fuchs & Michael Trauner

6. Department of Medicine, University Medical Centre Hamburg-Eppendorf, Hamburg, Germany

Gustav Buescher & Ansgar W. Lohse

7. Transplant and Regenerative Medicine Laboratory, Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK

Lu Wang, Chelsea Griffiths, Marnie L. Brown, William E. Scott III & Andrew J. Fisher

8. Department of Medicine, University of Cambridge, Cambridge, UK

William T. H. Gelson, Johannes Bargehr, Ravindra K. Gupta, Nicholas J. Matheson & Fotios Sampaziotis

9. Wellcome Sanger Institute, Hinxton, UK

Daniele Muraro & Ludovic Vallier

10. Centre of Excellence in Long-acting Therapeutics (CELT), Department of Pharmacology and Therapeutics, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK

Jo Sharp, Megan Neary, Helen Box, Lee Tatham, Paul Curley, Henry Pertinez, Jo Herriott, Edyta Kijak, Helen Cox, Chloe Bramwell, Anthony Valentijn & Andrew Owen

11. Department of Infection Biology and Microbiomes, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, UK

James Stewart

12. Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK

Mahnaz Darvish-Damavandi & Simon J. A. Buczacki

13. Academic Department of Medical Genetics, University of Cambridge, Cambridge, UK

Victoria L. Mulcahy & George F. Mells

14. Experimental Medicine and Immunotherapeutics, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK

Rhoda E. Kuc, Thomas L. Williams & Anthony P. Davenport

15. Department of Surgery, University of Cambridge and NIHR Cambridge Biomedical Research Centre, Cambridge, UK

Olivia C. Tysoe, Corrina Fear, Lisa Swift, Kourosh Saeb-Parsy, Paul Gibbs, Andrew J. Butler & Christopher J. E. Watson

16. Division of Cardiovascular Medicine, University of Cambridge, Cambridge, UK

Johannes Bargehr & Sanjay Sinha

17. Department of Radiology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

Sara S. Upponi

18. Roy Calne Transplant Unit, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

Kourosh Saeb-Parsy, Paul Gibbs, Andrew J. Butler & Christopher J. E. Watson

19. Department of Histopathology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

Susan E. Davies

20. Department of Medicine, Huddinge, Karolinska Institutet, Stockholm, Sweden

Axel Wester & Hannes Hagström

21. Norwegian PSC Research Center, Department of Transplantation Medicine, Division of Surgery, Inflammatory Diseases and Transplantation, Oslo University Hospital, Rikshospitalet, Oslo, Norway

Espen Melum

22. Research Institute of Internal Medicine, Division of Surgery, Inflammatory Diseases and Transplantation, Oslo University Hospital, Rikshospitalet, Oslo, Norway

Espen Melum

23. Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

Espen Melum

24. Section of Gastroenterology, Department of Transplantation Medicine, Division of Surgery, Inflammatory Diseases and Transplantation, Oslo University Hospital, Rikshospitalet, Oslo, Norway

Espen Melum

25. Hybrid Technology Hub Centre of Excellence, Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, Oslo, Norway

Espen Melum

26. MRC–University of Glasgow Centre for Virus Research, Glasgow, UK

Christopher J. R. Illingworth

27. Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, UK

Christopher J. R. Illingworth

28. Department of Health Behavior and Policy, Virginia Commonwealth University, Richmond, VA, USA

Bassam Dahman & Dustin R. Bastaich

29. Oxford Liver Unit, Translational Gastroenterology Unit, Oxford University Hospitals NHS Foundation Trust, University of Oxford, Oxford, UK

Thomas Marjot & Eleanor Barnes

30. Division of Gastroenterology and Hepatology, University of North Carolina, Chapel Hill, NC, USA

Andrew M. Moon, Alfred S. Barritt IV & Nicholas J. Matheson

31. Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

Gareth Corbett

32. Department of Histology and Embryology, School of Medicine, National and Kapodistrian University of Athens, Athens, Greece

Vassilis G. Gorgoulis

33. Ninewells Hospital and Medical School, University of Dundee, Dundee, UK

Vassilis G. Gorgoulis

34. Biomedical Research Foundation, Academy of Athens, Athens, Greece

Vassilis G. Gorgoulis

35. Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK

Joo-Hyeon Lee

36. NHS Blood and Transplant, Cambridge, UK

Nicholas J. Matheson

37. National Institute of Health Research (NIHR) Cambridge Biomedical Research Centre, and the NIHR Blood and Transplant Research Unit (BTRU) at the University of Cambridge in collaboration with Newcastle University and in partnership with NHS Blood and Transplant (NHSBT), Cambridge, UK

Christopher J. E. Watson

38. Berlin Institute of Health (BIH), BIH Centre for Regenerative Therapies (BCRT), Charité—Universitätsmedizin Berlin, Berlin, Germany

Ludovic Vallier

39. Max Planck Institute for Molecular Genetics, Berlin, Germany

Ludovic Vallier

Consortia

UK-PBC Consortium

- George F. Mells

Contributions

T.B. conceived and designed the study, performed experiments, acquired, interpreted and analysed the data, developed and validated the protocols described, generated the figures and wrote and edited the manuscript. M.M. performed SARS-CoV-2 infection experiments, collected and processed infectious samples and contributed to data interpretation and revision of the manuscript for intellectual content. G.J.W. and B.V.J. acquired, analysed and interpreted patients' data from the COVID-Hep and SECURE-Liver registries and contributed to revision of the manuscript for intellectual content. C.D.F. performed mouse experiments. G.B. contributed to the healthy volunteer study. L.W., C.G., M.L.B. and W.E.S. established lungs on ESNP and performed the lung ESNP experiment. P.P.-G. performed SARS-CoV-2 infection experiments in HEK293T cells and generated the luminescent biosensor. W.T.H.G. contributed to sample collection and revision of the manuscript for intellectual content. S.B. contributed to the generation of the luciferase reporter. S.D. performed transmission electron microscopy sample processing and preparation. D.M. performed bioinformatic analyses. J. Sharp performed hamster experiments and contributed to revision of the manuscript for intellectual content. M.N., H.B., L.T., J. Herriott, H.C., C.B., A.V., J. Stewart and E.K. performed

hamster experiments. P.C. and H.P. performed UDCA quantification in vivo. S.F. and P.M. contributed to viral infection experiments. S.S.V., M.D.-D. and T.W.M.C. expanded and provided cell lines, contributing to cell culture. V.L.M. contributed to blood sample collection. R.E.K. and T.L.W. provided primary tissue samples. J.A.H. and D.R. contributed to sample processing. V.G., M.V.-G. and O.C.T. provided samples for qPCR. J.B. provided primary tissue. S.S. and S.S.U. contributed to revision of the manuscript for intellectual content. L.S. and C.F. provided perfusionist support for ESNP experiments. K.S.-P. provided primary tissue and established livers on ESNP. S.E.D. provided primary tissue and contributed to histological analyses. D.C. and P.H. contributed to confocal imaging. A.W., H.H., E.M. and C.J.R.I. contributed to revision of the manuscript for intellectual content. B.D., D.R.B. and R.D.F. acquired and analysed patient data included in the VOCAL registry. T.M., E.B., A.M.M. and A.S.B. acquired and analysed patient data included in the COVID-Hep and SECURE-Liver registries. R.K.G., S.B. and V.G.G. contributed to revision of the manuscript for intellectual content. G.C. and A.P.D. provided primary tissue samples and contributed to revision of the manuscript for intellectual content. S.J.A.B. and J.-H.L. provided primary organoids and contributed to data interpretation and revision of the manuscript for intellectual content. N.J.M. supervised the SARS-CoV-2 infection experiments in HEK293T cells, contributed to data interpretation and reviewed the manuscript for intellectual content. M.T. provided animals and reviewed the manuscript for intellectual content. A.J.F. provided donor lungs and co-designed the ESNP lung experiment, contributed to ESNP experiments and reviewed the manuscript for intellectual content. P.G. provided tissue and bile samples and reviewed the manuscript for intellectual content. A.J.B. and C.J.E.W. provided tissue and bile samples, contributed to ESNP experiments and reviewed the manuscript for intellectual content. G.F.M. provided samples and contributed to revision of the manuscript for intellectual content and sample provision. G.D. contributed to revision of the manuscript for intellectual content. A.O. led the hamster study, contributed to data interpretation and reviewed the manuscript for intellectual content. A.W.L. acquired and analysed patient data included in the COVID-Hep and SECURE-Liver registries, performed the volunteer study and contributed to revision of the manuscript for intellectual content. L.V. designed and conceived the study, interpreted the data and edited the manuscript. F.S.

designed and conceived the study, interpreted the data, edited the manuscript and performed ESNP experiments. All of the authors approved the manuscript.

Corresponding authors

Correspondence to Teresa Brevini, Ludovic Vallier or Fotios Sampaziotis.

Ethics declarations

Competing interests

F.S., L.V. and K.S.-P. are founders and shareholders of Bilitech. L.V. is a founder and shareholder of DEFINIGEN. The remaining authors declare no competing interests.

Peer review

Peer review information

Nature thanks Wendong Huang, David Jones, Stanley Perlman, Frank Tacke and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Expression of SARS-CoV-2 entry genes in cholangiocytes.

(**a**) Schematic illustration of different primary human cholangiocyte populations corresponding to different areas of the biliary tree and COs derived from different areas of the biliary tree grown in absence or presence of the bile acids. (**b**) UMAP plot illustrating different cholangiocyte populations from (a) analysed by scRNA-seq. (**c**) UMAP plots showing that viral entry related genes are predominantly expressed in extrahepatic cholangiocytes and COs treated with bile acids. (**d–e**) Immunofluorescence illustrating that ACE2 is expressed in extrahepatic cholangiocytes (**d**) and that SARS-CoV-2 infects gall bladder cholangiocytes of patients with COVID-19 but not intrahepatic cholangiocytes (**e**). N = 4 independent samples. Scale bars 50 µm. (**f**) QPCR confirming detection of SARS-CoV-2 RNA in bile of patients with COVID-19. Housekeeping gene, HMBS; n = 4; two-tailed Mann–Whitney test; centre line, median; box, interquartile range (IQR); whiskers, range; bars, standard deviation. (**g**) Violin plot of scRNA-seq data from (b) showing that COs upregulate ACE2 when treated with bile acids regardless of their region of origin. (**h**) QPCR validating that upon treatment with the bile acid CDCA COs assume a gall bladder identity expressing the gall bladder marker SOX17 and upregulating ACE2 at levels comparable to primary gall bladder. Housekeeping gene, HMBS; n = 4 independent experiments; one-way ANOVA adjusted for multiple comparisons; ns, non-significant; centre line, median; box, interquartile range (IQR); whiskers, range; bars, standard deviation. (**i**) Immunofluorescence showing that CDCA induces ACE2 expression in GCOs. N = 4 independent experiments. Scale bars 50 µm.

Extended Data Fig. 2 CDCA-treated GCOs can be infected with SARS-CoV-2.

(**a**) Schematic representation of the methodology used to infect GCOs with SARS-CoV-2 and test the capacity of SARS-CoV-2 virions produced in GCOs to infect new (uninfected) cells. (**b**) Immunofluorescence validating SARS-CoV-2 infection in GCOs. N = 4 independent experiments. Scale bars 50 µm. (**c**) QPCR confirming infection of CDCA-treated GCOs with SARS-CoV-2 propagated in VERO E6 cells (top panel) and with SARS-CoV-2 propagated in GCOs treated with CDCA (bottom panel), illustrating that SARS-CoV-2 produced in GCOs+CDCA retains its infectious capacity. Scale bars 50 µm; housekeeping gene, GAPDH; n = 4 independent

experiments; Kruskal–Wallis test adjusted for multiple comparisons; centre line, median; box, interquartile range (IQR); whiskers, range; bars, standard deviation. (**d**) QPCR showing upregulation of innate immune and antiviral response genes in GCOs+CDCA following SARS-CoV-2 infection. Housekeeping gene, GAPDH; n = 4, 2 biological and 2 technical replicates; two-tailed Mann–Whitney test (IL-1β, IL-6, IFNα) and two-tailed unpaired t-test (TNF, IFNλ). (CDCA concentration, 10 μM); centre line, median; box, interquartile range (IQR); whiskers, range; bars, standard deviation. (**e**) Transmission electron micrograph of uninfected COs (left panel) and COs infected with SARS-CoV-2 in the absence (central panel) or presence of CDCA (right panel) showing key morphological features of viral infection and cell death, such as production of viral particles, formation of pathologic vacuoles (Vc) and swollen mitochondria (Mt). N = 3 independent experiments. Scale bars 5 μm.

Extended Data Fig. 3 FXR is present and active in tissues affected by COVID-19 and their corresponding CDCA-treated organoids.

(**a–b**) Immunofluorescence images confirming expression of FXR in primary human gall bladder, lungs and intestinal tissue (**a**) and in corresponding primary organoids treated with physiological levels of bile acids (CDCA, 10 μM) (**b**). N = 4 independent experiments. White scale bars 100 μm; grey scale bar 25 μm. (**c**) QPCR analysis validating expression of FXR and its downstream effector SHP in primary tissue and corresponding organoids in presence or absence of physiological levels of bile acids (CDCA). CDCA treatment increases FXR and SHP expression in organoids to levels that are closer to primary tissue. Housekeeping gene, HMBS; n = 4 independent experiments; one-way ANOVA adjusted for multiple comparisons; centre line, median; box, interquartile range (IQR); whiskers, range; bars, standard deviation. AOs, airway organoids; IOs, intestinal organoids.

Extended Data Fig. 4 UDCA and ZGG require FXR to reduce ACE2 and SARS-CoV-2 infection.

(**a**–**b**) Immunofluorescence (**a**) and QPCR (**b**) showing downregulation of FXR expression following FXR knockdown (KD) in COs. Scale bars 25 µm. Housekeeping gene HMBS. One-way ANOVA adjusted for multiple comparisons; n = 4 independent experiments; centre line, median; box, interquartile range; whiskers, range; bars, standard deviation. (**c**) QPCR on FXR KD COs showing no change in the expression of ACE2 and the FXR downstream effector SHP following treatment with CDCA, UDCA or ZGG, demonstrating that FXR is indispensable for regulating ACE2 and SHP through these compounds. ACE2 and SHP expression in wild-type organoids shown in Fig. 1c and Extended Data Fig. 5a. Housekeeping gene, HMBS; n = 4; one-way ANOVA adjusted for multiple comparisons; ns, non-significant; centre line, median; box, interquartile range; whiskers, range; bars, standard deviation. (**d**) QPCR quantifying SARS-CoV-2 RNA 24 h post infection in FXR KD COs treated with CDCA, UDCA or ZGG. SARS-CoV-2 infection in the presence of FXR is shown in Fig. 1e. Housekeeping gene, GAPDH; n = 4 independent experiments; one-way ANOVA adjusted for multiple comparisons; centre line, median; box, interquartile range; whiskers, range; bars, standard deviation. (CDCA, UDCA and ZGG concentration, 10 µM). (**e**) Schematic illustrating the luciferase reporter construct containing the FXRE IR-1 identified in Fig. 1a and the mutagenesis strategy used in panel (f). (**f**) Luciferase reporter assay in HEK293 cells showing the transcriptional activity associated with the IR-1 located in the ACE2 promoter upon treatment with CDCA, UDCA or ZGG. Site-directed mutagenesis on IR-1 abolishes FXR binding/transactivation confirming the specificity of FXR binding on the ACE2 IR-1. IR-1 located in the SHP promoter used as positive control. N = 3 independent experiments; one-way ANOVA adjusted for multiple comparisons; bars, standard deviations. (CDCA, UDCA and ZGG concentration, 50 µM).

Extended Data Fig. 5 FXR modulation in biliary, airway and intestinal organoids.

(**a**) QPCR analysis in primary airway, biliary and intestinal organoids demonstrating that CDCA activates FXR, whereas UDCA and ZGG inhibit it, as evidenced by corresponding changes in the expression of the FXR downstream target SHP. Housekeeping gene, HMBS; n = 4 independent

experiments; one-way ANOVA adjusted for multiple comparisons; centre line, median; box, interquartile range (IQR); whiskers, range; bars, standard deviation. (**b**) Immunofluorescence showing ACE2 expression levels in primary organoids in absence of bile acids. The panel is complementary to Fig. 1d showing the modulation of ACE2 following FXR activation (CDCA) and inhibition (UDCA/ZGG). N = 4 independent experiments. Yellow scale bars 50 µm; grey scale bar 25 µm. (**c**) Flow cytometry histograms showing changes in ACE2 levels upon modulation of FXR activity in primary airway, biliary and intestinal organoids. n = 3 independent experiments. (CDCA, UDCA and ZGG concentration, 10 µM). (**d**) Dose–response curves showing the effect of 0.01 µM – 1 mM of CDCA, UDCA and ZGG on the expression of ACE2 in primary airway, biliary and intestinal organoids (n = 3 independent experiments). Response defined as percentage of the maximal ACE2 expression level for each condition via QPCR. Bars, SEM. (**e**) Percentage of non-viable cells following treatment of airway, biliary and intestinal organoids with CDCA, UDCA and ZGG at a range of 0.1 µM – 100 µM showing that these compounds cause minimal cell death within the tested range. N = 3 independent experiments. Bars, range. (**f**) Resazurin assay (PrestoBlue) showing that treatment with 10 µM of CDCA, UDCA or ZGG does not affect cellular viability. N = 4 independent experiments; one-way ANOVA adjusted for multiple comparisons; ns, non-significant; centre line, median; box, interquartile range (IQR); whiskers, range; bars, standard deviation.

Extended Data Fig. 6 ACE2 downregulation is required for the UDCA- or ZGG-mediated reduction in SARS-CoV-2 infection.

(**a**-**b**) QPCR analysis (**a**) and immunofluorescence (**b**) illustrating ACE2 and FXR expression in wild-type HEK293 cells and HEK293T cells stably expressing ACE2. Primary human airway tissue used as positive control. Housekeeping gene, HMBS; n = 4; one-way ANOVA adjusted for multiple comparisons; centre line, median; box, interquartile range (IQR); whiskers, range; bars, standard deviation. Scale bars 100 µm. (**c**) SARS-CoV-2 infection of HEK293T cells genetically engineered to stably express ACE2. Cells were treated with the indicated doses of CDCA, UDCA or ZGG, infected with SARS-CoV-2 at an MOI of 0.01, and analysed after 24 h. The SARS-CoV-2 RdRp inhibitor remdesivir and a neutralizing antibody

cocktail blocking the interaction between SARS-CoV-2 spike and ACE2 (REGN-COV2) were included as positive controls. N = 3; one-way ANOVA adjusted for multiple comparisons; mean values ± SEM.

Extended Data Fig. 7 FXR inhibition reduces ACE2 levels in vivo.

(**a**) Schematic representation of the experiment performed. (**b**) QPCR showing that treatment with UDCA in FVB/N mice reduces ACE2 levels in lung, gall bladder and intestinal tissue. Housekeeping gene, GAPDH; n = 4 animals per group (UDCA vs. no UDCA control group, see Methods); unpaired two-tailed t-test; centre line, median; box, interquartile range (IQR); whiskers, range; bars, standard deviation. (**c**) Immunofluorescence images showing ACE2 levels upon treatment with UDCA in respiratory, biliary and intestinal epithelium in FVB/N mice. N = 4 mice per group. White scale bars 100 µm; yellow scale bars 50 µm. (**d**) QPCR showing that treatment with UDCA in Syrian golden hamsters reduces ACE2 levels in the gall bladder and intestinal tissue. Housekeeping gene, GAPDH; n = 5 vehicle/No UDCA group vs n = 3 UDCA group; unpaired two-tailed t-test; centre line, median; box, interquartile range (IQR); whiskers, range; bars, standard deviation. (**e**) Immunofluorescence images showing ACE2 levels upon treatment with UDCA in biliary and intestinal epithelium in Syrian golden hamsters. N = 3 hamsters per group. White scale bars 100 µm.

Source data

Extended Data Fig. 8 UDCA plasma concentration in vivo.

(**a**) UDCA concentration in the plasma of hamsters over 7 days of treatment with 416 mg/kg/day of UDCA. N = 6 animals/group; line, median. (**b**) Viral titre showing levels of infectious virus as measured by plaque assay in lungs in directly inoculated hamsters and sentinel animals treated with UDCA or vehicle and co-housed with infected animals. Samples were collected after 4 days of co-housing. N = 6 UDCA vs n = 3 vehicle animals; Kruskal–Wallis test with Dunn’s correction for multiple comparisons. Animals from each experiment are represented with different symbols; line, median.

Source data

Extended Data Fig. 9 FXR inhibition reduces SARS-CoV-2 infection in human organs ex vivo.

(**a**) Photograph and schematic of the liver ex situ normothermic perfusion (ESNP) experiment performed; including type of samples and timeline. 0 h: 0 h baseline sample collection and UDCA/carrier administration. The 0 h samples were collected immediately prior to UDCA administration. 12 h: 12 h after UDCA/carrier administration. For each time point, 4 independent tissue samples were obtained from the grafts' gall bladder and used for ACE2 measurement and viral infection (n = 4 gall bladder tissue samples per time point). (**b**) ACE2 enzymatic activity measurement showing that ESNP with UDCA reduces ACE2 activity in the circulating perfusate, compared to carrier only control. N = 4 independent samples; unpaired two-tailed t-test; centre line, median; box, interquartile range (IQR); whiskers, range; bars, standard deviation. (**c**) Immunofluorescence images showing ACE2 expression in gall bladder cholangiocytes and cells of the vasculature (smooth muscle and endothelial cells) in human livers before and after ESNP with UDCA/carrier. N = 4 independent samples. White scale bars 100 µm; yellow scale bars 50 µm, grey scale bars 25 µm. (**d**) QPCR demonstrating that ESNP with UDCA reduces ACE2 levels in gall bladder cholangiocytes, compared to carrier only control. Housekeeping gene, HMBS; n = 4; unpaired two-tailed t-test; centre line, median; box, interquartile range (IQR); whiskers, range; bars, standard deviation. (**e-f**) QPCR (**e**) and immunofluorescence (**f**) showing that UDCA reduces SARS-CoV-2 infection in human gall bladder ex vivo. Housekeeping gene, GAPDH; n = 4; one-way ANOVA adjusted for multiple comparisons; centre line, median; box, interquartile range (IQR); whiskers, range; bars, standard deviation. Scale bars 25 µm. (UDCA, concentration, 2,000 ng/ml).

Extended Data Fig. 10 UDCA is associated with lower levels of ACE2 in patients with PBC.

(**a**) Schematic illustrating the patient cohorts compared in (b). (**b**) Multiple linear regression analysis of serum ACE2 demonstrates that UDCA

correlates with lower ACE2 levels in patients with primary biliary cholangitis (PBC) receiving UDCA (n = 308) vs. patients with PBC who were naïve to treatment (n = 62). Values plotted are β coefficients. (BMI, body mass index; ALP, alkaline phosphatase). Bars, 95% CI.

Supplementary information

Supplementary Information

This file contains Supplementary Fig. 1 (Gating strategy for flow cytometry analyses) and Supplementary Tables 1–5 and 7–9 (see separate Excel file for Supplementary Table 6).

Reporting Summary

Supplementary Table 6

UK-PBC cohort characteristics relative to Extended Data Fig. 10.

Supplementary Data

This file contains code for Extended Data Fig. 10.

Source data

Source Data Fig. 2

Source Data Extended Data Fig. 7

Source Data Extended Data Fig. 8

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing,

adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Brevini, T., Maes, M., Webb, G.J. *et al.* FXR inhibition may protect from SARS-CoV-2 infection by reducing ACE2. *Nature* **615**, 134–142 (2023). https://doi.org/10.1038/s41586-022-05594-0

- Received: 03 May 2021
- Accepted: 23 November 2022
- Published: 05 December 2022
- Issue Date: 02 March 2023
- DOI: https://doi.org/10.1038/s41586-022-05594-0

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

This article is cited by

- **Closing the door to SARS-CoV-2**

    - Sarah Crunkhorn

    Nature Reviews Drug Discovery (2023)

- **Actionable targets to reduce COVID-19 severity**

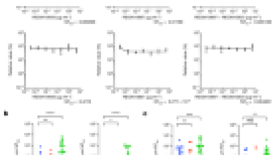
    - Shiu Lun Au Yeung
    - Shan Luo
    - Kin On Kwok

    Nature Metabolism (2023)

A liver drug reduces SARS-CoV-2 entry into cells

Nature Research Briefing 05 Dec 2022

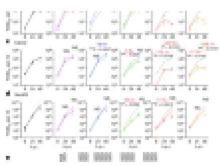
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-022-05594-0



Altered TMPRSS2 usage by SARS-CoV-2 Omicron impacts infectivity and fusogenicity

01 February 2022

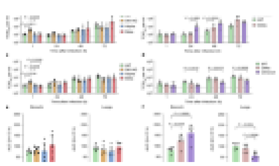
Bo Meng, Adam Abdullahi, … Ravindra K. Gupta



Attenuated replication and pathogenicity of SARS-CoV-2 B.1.1.529 Omicron

21 January 2022

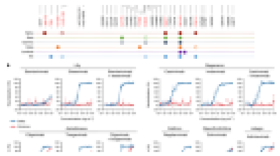
Huiping Shuai, Jasper Fuk-Woo Chan, … Hin Chu



SARS-CoV-2 Omicron variant replication in human bronchus and lung ex vivo

01 February 2022

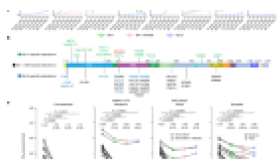
Kenrie P. Y. Hui, John C. W. Ho, … Michael C. W. Chan



[Considerable escape of SARS-CoV-2 Omicron to antibody neutralization]

23 December 2021

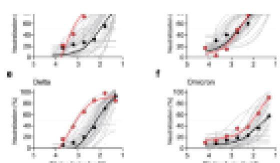
Delphine Planas, Nell Saunders, … Olivier Schwartz



[Antibody evasion properties of SARS-CoV-2 Omicron sublineages]

03 March 2022

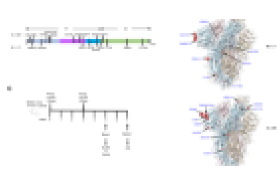
Sho Iketani, Lihong Liu, … David D. Ho



[A broadly neutralizing antibody protects Syrian hamsters against SARS-CoV-2 Omicron challenge]

23 June 2022

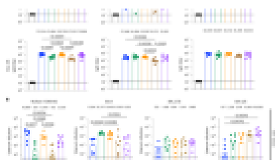
Biao Zhou, Runhong Zhou, … Zhiwei Chen



Naturally occurring spike mutations influence the infectivity and immunogenicity of SARS-CoV-2

12 October 2022

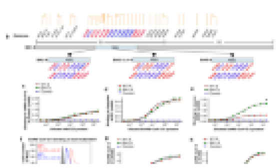
Qiaoli Peng, Runhong Zhou, … Zhiwei Chen



Bivalent SARS-CoV-2 mRNA vaccines increase breadth of neutralization and protect against the BA.5 Omicron variant in mice

20 October 2022

Suzanne M. Scheaffer, Diana Lee, … Michael S. Diamond



Effective vaccination strategy using SARS-CoV-2 spike cocktail against Omicron and other variants of concern

19 December 2022

Juan Shi, Gang Wang, … Lanying Du

- Article
- Published: 11 January 2023

Spike and nsp6 are key determinants of SARS-CoV-2 Omicron BA.1 attenuation

- Da-Yuan Chen,
- Chue Vin Chin,
- Devin Kenney,
- Alexander H. Tavares,
- Nazimuddin Khan,
- Hasahn L. Conway,
- GuanQun Liu,
- Manish C. Choudhary,
- Hans P. Gertje,
- Aoife K. O'Connell,
- Scott Adams,
- Darrell N. Kotton,
- Alexandra Herrmann,
- Armin Ensser,
- John H. Connor,
- Markus Bosmann,
- Jonathan Z. Li,
- Michaela U. Gack,
- Susan C. Baker,
- Robert N. Kirchdoerfer,
- Yachana Kataria,
- Nicholas A. Crossland,
- Florian Douam &
- …
- Mohsan Saeed

Show authors

Nature **volume 615**, pages 143–150 (2023)

- 12k Accesses
- 1 Citations
- 289 Altmetric
- Metrics details

Abstract

The SARS-CoV-2 Omicron variant is more immune evasive and less virulent than other major viral variants that have so far been recognized[1,2,3,4,5,6,7,8,9,10,11,12]. The Omicron spike (S) protein, which has an unusually large number of mutations, is considered to be the main driver of these phenotypes. Here we generated chimeric recombinant SARS-CoV-2 encoding the *S* gene of Omicron (BA.1 lineage) in the backbone of an ancestral SARS-CoV-2 isolate, and compared this virus with the naturally circulating Omicron variant. The Omicron S-bearing virus robustly escaped vaccine-induced humoral immunity, mainly owing to mutations in the receptor-binding motif; however, unlike naturally occurring Omicron, it efficiently replicated in cell lines and primary-like distal lung cells. Similarly, in K18-hACE2 mice, although virus bearing Omicron S caused less severe disease than the ancestral virus, its virulence was not attenuated to the level of Omicron. Further investigation showed that mutating non-structural protein 6 (nsp6) in addition to the S protein was sufficient to recapitulate the attenuated phenotype of Omicron. This indicates that although the vaccine escape of Omicron is driven by mutations in S, the pathogenicity of Omicron is determined by mutations both in and outside of the S protein.

Main

As of December 2022, the successive waves of the COVID-19 pandemic have been driven by five major SARS-CoV-2 variants, known as variants of concern (VOC): Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2 and AY lineages) and Omicron (BA lineages)[13]. Omicron is the most recently recognized VOC, and was first documented in South Africa,

Botswana and in a traveller from South Africa in Hong Kong in November 2021 (GISAID ID: EPI_ISL_7605742)[14,15]. It quickly swept through the world, displacing the previously dominant Delta variant within weeks and accounting for most new SARS-CoV-2 infections by January 2022 (refs. [16,17,18]). At least five lineages of Omicron have so far been identified: BA.1, BA.2, BA.3, BA.4 and BA.5. BA.1 (hereafter referred to as Omicron) exhibits a remarkable escape from infection- and vaccine-induced humoral immunity[3,19]. Furthermore, it is less virulent than other VOCs in humans and in vivo models of infection[4,5,7,11,12,20]. Omicron differs from the prototype SARS-CoV-2 isolate, Wuhan-Hu-1, by 59 amino acids; 37 of these changes are in the S protein, raising the possibility that S is at the heart of Omicron's pathogenic and antigenic behaviour.

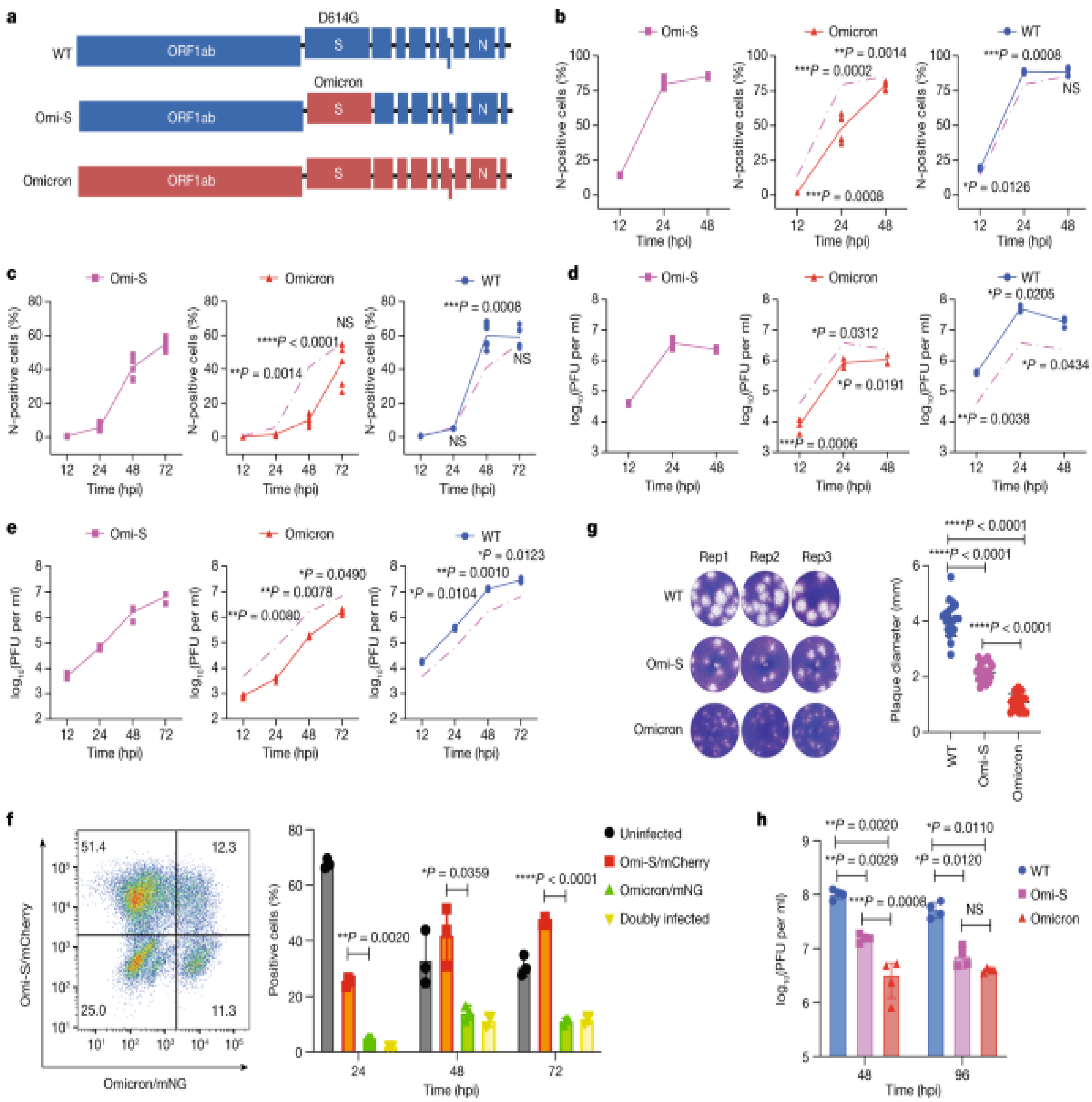
S mutations affect Omicron replication in cell culture

The Omicron S protein contains 30 amino acid substitutions, 6 deletions and one insertion of 3 amino acids in length, as compared to Wuhan-Hu-1 (Extended Data Fig. 1). Twenty-five of these changes are unique to Omicron relative to other VOCs, although some of them have been reported in waste water and minor SARS-CoV-2 variants[21,22]. To test the role of the S protein in Omicron phenotype, we generated a chimeric recombinant virus containing the *S* gene of Omicron (USA-lh01/2021) and all other genes of an ancestral SARS-CoV-2 (GISAID EPI_ISL_2732373)[23] (Extended Data Fig. 2a). This chimeric virus, named Omi-S, was made by using a modified form of circular polymerase extension reaction (CPER)[24] (Extended Data Fig. 2b) that yielded highly concentrated virus stocks, containing 0.5×10^6–5×10^6 plaque-forming units (PFU) per ml, from transfected cells within two days of transfection (Extended Data Fig. 2c,d), obviating the need for further viral amplification.

We first compared the infection efficiency of Omi-S with an ancestral D614G-containing virus (GISAID EPI_ISL_2732373; generated by CPER; hereafter referred to as wild type (WT)) and an Omicron isolate (USA-lh01/2021) in cell culture (Fig. 1a). For this, we infected ACE2- and TMPRSS2-expressing Caco-2 (hereafter, ACE2/TMPRSS2/Caco-2) and

Vero E6 cells with Omi-S, WT and Omicron at a multiplicity of infection (MOI) of 0.01, and monitored viral propagation by flow cytometry and plaque formation assay. The WT virus and Omi-S spread rapidly in ACE2/TMPRSS2/Caco-2 cells, yielding 89% and 80% infected cells, respectively, at 24 hours post-infection (hpi) (Fig. 1b). By contrast, Omicron replicated more slowly, leading to 48% infected cells at 24 hpi. A similar pattern was seen in Vero E6 cells, in which 60% and 41% of cells were positive for WT and Omi-S, respectively, at 48 hpi, as opposed to 10% for Omicron (Fig. 1c). The plaque assay showed that although both Omi-S and Omicron produced lower levels of infectious virus particles compared with WT, the viral titre of Omi-S was significantly higher than that of Omicron. In ACE2/TMPRSS2/Caco-2 cells, Omi-S produced 5.1-fold ($P = 0.0006$) and 5.5-fold ($P = 0.0312$) more infectious particles than Omicron at 12 hpi and 24 hpi, respectively (Fig. 1d). Similarly, in Vero E6 cells, the infectious virus titres of Omi-S were 17-fold ($P = 0.0080$) and 11-fold ($P = 0.0078$) higher than that of Omicron at 24 hpi and 48 hpi, respectively (Fig. 1e). The difference between viruses became less obvious at later time points owing to higher cytotoxicity caused by Omi-S compared to Omicron (Extended Data Fig. 3a).

Fig. 1: Effect of S on the in vitro growth kinetics of Omicron.



a, Schematic of viruses. S, spike; N, nucleocapsid. **b**–**e**, ACE2/TMPRSS2/Caco-2 cells (**b**,**d**) and Vero E6 cells (**c**,**e**) were infected at an MOI of 0.01, and the percentage of N-positive cells ($n = 6$ replicates) (**b**,**c**) and the release of infectious particles ($n = 3$ replicates) (**d**,**e**) were determined by flow cytometry and by plaque assay, respectively. **f**, ACE2/TMPRSS2/Caco-2 cells were infected with virus mixtures at a 1:1 ratio to obtain the final MOI of 0.005 for each virus. The cells were fixed at the indicated times and subjected to flow cytometry. Left, representative dot plot; right, percentage of uninfected, Omi-S/mCherry-infected, Omicron/mNeoGreen (Omicron/mNG)-infected and doubly infected cells. Singly infected cells were used for compensation. Error bars, mean ± s.d. (*n*

= 3 replicates). **g**, Plaque sizes. Left, representative images of plaques on ACE2/TMPRSS2/Caco-2 cells. Right, the diameter of plaques is plotted as the mean ± s.d. of 20 plaques per virus. **h**, Human iAT2 epithelial cells were infected at an MOI of 2.5 for 48 h or 96 h. The apical side of cells was washed with 1× phosphate-buffered saline (PBS) and the levels of infectious virus particle were measured by plaque assay. Error bars, mean ± s.d. ($n = 4$ replicates). Experiments were repeated twice, with each experimental repeat containing 3 (**b**–**g**) or 4 (**h**) replicates. P values were calculated by a two-tailed, unpaired t-test with Welch's correction. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ and ****$P < 0.0001$; NS, not significant. The gating strategy for flow cytometry is shown in Supplementary Fig. 1.

The increased replication efficiency of Omi-S relative to Omicron was preserved when tested at varying MOIs (Extended Data Fig. 3b). We further confirmed the fitness advantage of Omi-S over Omicron by a direct competition assay. For this, we first generated recombinant Omicron (rOmicron), which, in our cell culture assays, mimicked the replication kinetics of natural Omicron (Extended Data Fig. 4). Next, we created mCherry-containing Omi-S and mNeonGreen-containing Omicron, and inoculated ACE2/TMPRSS2/Caco-2 cells with these viruses mixed at a 1:1 ratio. Flow cytometric analysis of infected cells at various times of infection showed that Omi-S/mCherry was clearly superior to Omicron/mNeonGreen in terms of replication (Fig. 1f). Finally, the higher infection efficiency of Omi-S was also reflected in the plaque size; although WT virus produced the largest plaques (around 4.1 mm), the size of Omi-S plaques (around 2.2 mm) was twofold ($P < 0.0001$) larger than that of Omicron plaques (around 1.1 mm) (Fig. 1g). These results indicate that although mutations in the S protein influence the infection efficiency of Omicron, they do not fully explain the Omicron phenotype.

Several lines of evidence indicated that the S protein incorporated into Omi-S behaved the same way as in natural Omicron. For instance, as described previously[20,25], Omicron S was poorly cleaved compared to WT S; whereas 71% of S in WT virions was in the cleaved form, only 45% and 47% was cleaved in Omi-S and Omicron, respectively (Extended Data Fig. 5a). The same pattern of S cleavage was evident in virus-infected cells (WT, 63% cleaved; Omi-S, 33% cleaved; Omicron, 42% cleaved) (Extended Data Fig. 5b). These experiments also revealed that Omicron S was inefficiently

incorporated into virus particles compared to WT S (S to nucleocapsid (N) ratio: 3.40 for WT virus, 1.91 for Omi-S and 2.04 for Omicron) (Extended Data Fig. 5a). Similarly, both Omi-S and Omicron produced smaller syncytia compared to the WT virus, an observation that has previously been reported for Omicron[20,26] (Extended Data Fig. 5c). Finally, consistent with the published literature[25], Omi-S and Omicron showed a preference for cathepsin-mediated entry, as reflected by their higher sensitivity to the cathepsin inhibitor E64d (Extended Data Fig. 6).

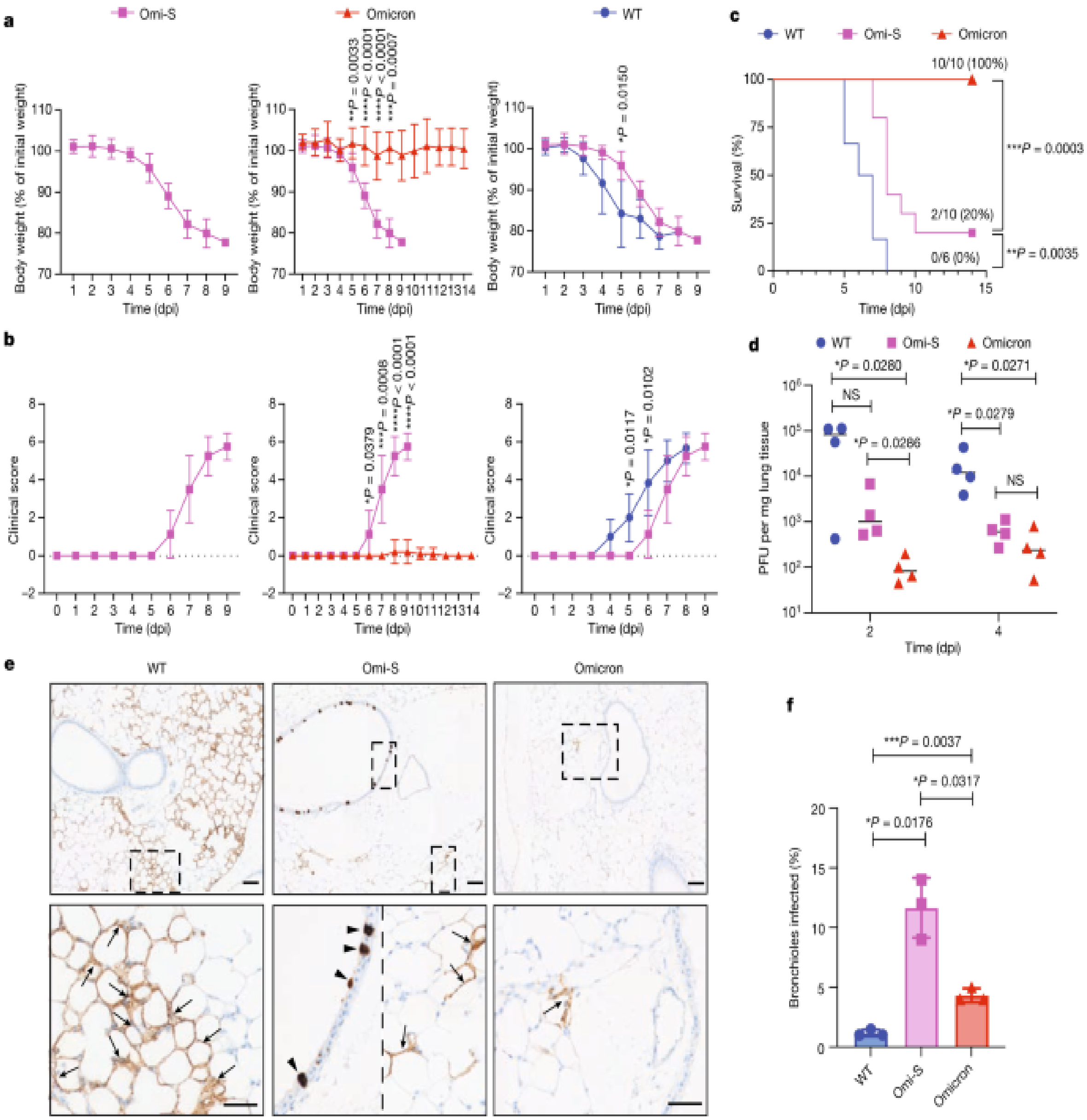
We next compared the replication kinetics of WT, Omi-S and Omicron in lung epithelial cells, which form a major viral replication site in patients with COVID-19 (refs. [27,28]). Accordingly, we used human induced pluripotent stem (iPS)-cell-derived lung alveolar type 2 epithelial (iAT2) cells. AT2 cells represent an essential cell population in the distal lung and constitute one of the primary targets of SARS-CoV-2 infection[28,29]. We infected iAT2 cells, grown as an air–liquid interface (ALI) culture, at an MOI of 2.5 and monitored the secretion of viral progeny on the apical side of cells at 48 hpi and 96 hpi. In congruence with the results obtained from cell lines, WT virus produced the highest levels of infectious virus particles (Fig. 1h). Among Omi-S and Omicron, the former yielded around fivefold ($P = 0.0008$) higher infectious viral titre at 48 hpi. The viral titres for WT and Omi-S decreased at 96 hpi compared to 48 hpi owing to the cytopathic effect (CPE) of infection. However, no CPE was seen for Omicron, leading to sustained production of infectious virions. Overall, these results corroborate the conclusion that mutations in S do not fully account for the attenuated replication capacity of Omicron in cell culture.

Minimal role of S in Omicron pathogenicity in mice

To examine whether Omi-S exhibits higher in vivo fitness compared with Omicron, we investigated the infection outcome of Omi-S relative to WT SARS-CoV-2 and Omicron in K18-hACE2 mice. In agreement with the published literature[4,5], intranasal inoculation of mice (aged 12–20 weeks) with Omicron (10^4 PFU per mouse) caused no significant weight loss, whereas inoculation with WT virus triggered a rapid decrease in body

weight, with all mice losing over 20% of their initial body weight by 8 days post-infection (dpi) (Fig. 2a). Notably, 80% of mice infected with Omi-S also lost over 20% of their body weight by 9 dpi (Fig. 2a and Extended Data Fig. 7a). The evaluation of clinical scores (a cumulative measure of weight loss, abnormal respiration, aberrant appearance, reduced responsiveness and altered behaviour) also revealed a similar pattern; whereas Omicron-infected mice exhibited few to no signs of clinical illness, the health of those infected with WT and Omi-S rapidly deteriorated, with the WT virus causing a more severe disease ($P = 0.0102$) (Fig. 2b and Extended Data Fig. 7b). Because SARS-CoV-2 causes fatal infection in K18-hACE2 mice[4], we compared the survival of mice after viral infection. Agreeing with the results of body-weight loss and clinical score, WT and Omi-S caused mortality rates of 100% (6/6) and 80% (8/10), respectively. By contrast, all mice infected with Omicron survived (Fig. 2c). These findings, which are consistent with a recent publication[30], indicate that the S protein is not the exclusive determinant of Omicron's pathogenicity in K18-hACE2 mice.

Fig. 2: Role of S in Omicron pathogenicity.



a–**c**, Male and female K18-hACE2 mice (aged 12–20 weeks) were intranasally inoculated with 1 × 10^4 PFU of WT (*n* = 6 mice), Omi-S (*n* = 10 mice) or Omicron (*n* = 10 mice) virus. Two independently generated virus stocks were used in this experiment. Body weight (**a**), clinical score (**b**) and survival (**c**) were monitored daily for 14 days. Mice that lost 20% of their initial body weight were euthanized. **d**,**e**, K18-hACE2 mice were intranasally inoculated with 1 × 10^4 PFU of WT (*n* = 14 mice), Omi-S (*n* = 14 mice) and Omicron (*n* = 14 mice). Lung samples of the infected mice were collected at 2 or 4 dpi to determine the viral titre (*n* = 4 mice) (**d**) or for immunohistochemistry (IHC) detection of the N protein (*n* = 3 mice) (**e**). In

e, representative IHC images showing SARS-CoV-2 N (brown colour) in alveoli (arrows) and bronchioles (arrowheads) in mice lungs at 2 dpi are presented. Scale bars, 100 µm. **f**, Percentage of N-positive bronchioles in the lungs of infected mice ($n = 3$ mice) at 2 dpi. Each dot represents an infected mouse. Statistical significance was determined using a two-tailed, unpaired *t*-test with Welch's correction (**a**,**b**,**d**,**f**) and log-rank (Mantel-Cox) test (**c**). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ and $^{****}P < 0.0001$; NS, not significant.

Source data

Next, we compared the propagation of Omi-S with Omicron and WT SARS-CoV-2 in the lungs and nasal turbinates of K18-hACE2 mice. The mice (12–20 weeks old) were intranasally challenged with 10^4 PFU (seven mice per virus), and viral titres in mice lungs were measured at 2 and 4 dpi. Consistent with in vitro findings, the infectious virus titre in the lungs of WT-infected mice was higher than that detected in mice infected with the other two viruses (Fig. 2d). Notably, however, Omi-S-infected mice produced 30-fold ($P = 0.0286$) more infectious virus particles compared with Omicron-infected mice at 2 dpi. The titre decreased at 4 dpi for WT- and Omi-S-infected mice, but it showed an increasing trend for Omicron-infected mice, pointing to the possibility of mild but persistent infection by Omicron in K18-hACE2 mice. All three variants recovered from the lungs of mice maintained the same plaque size phenotype as the original inoculum, indicating that replication in mice had no detectable effect on genotypes of these viruses (data not shown).

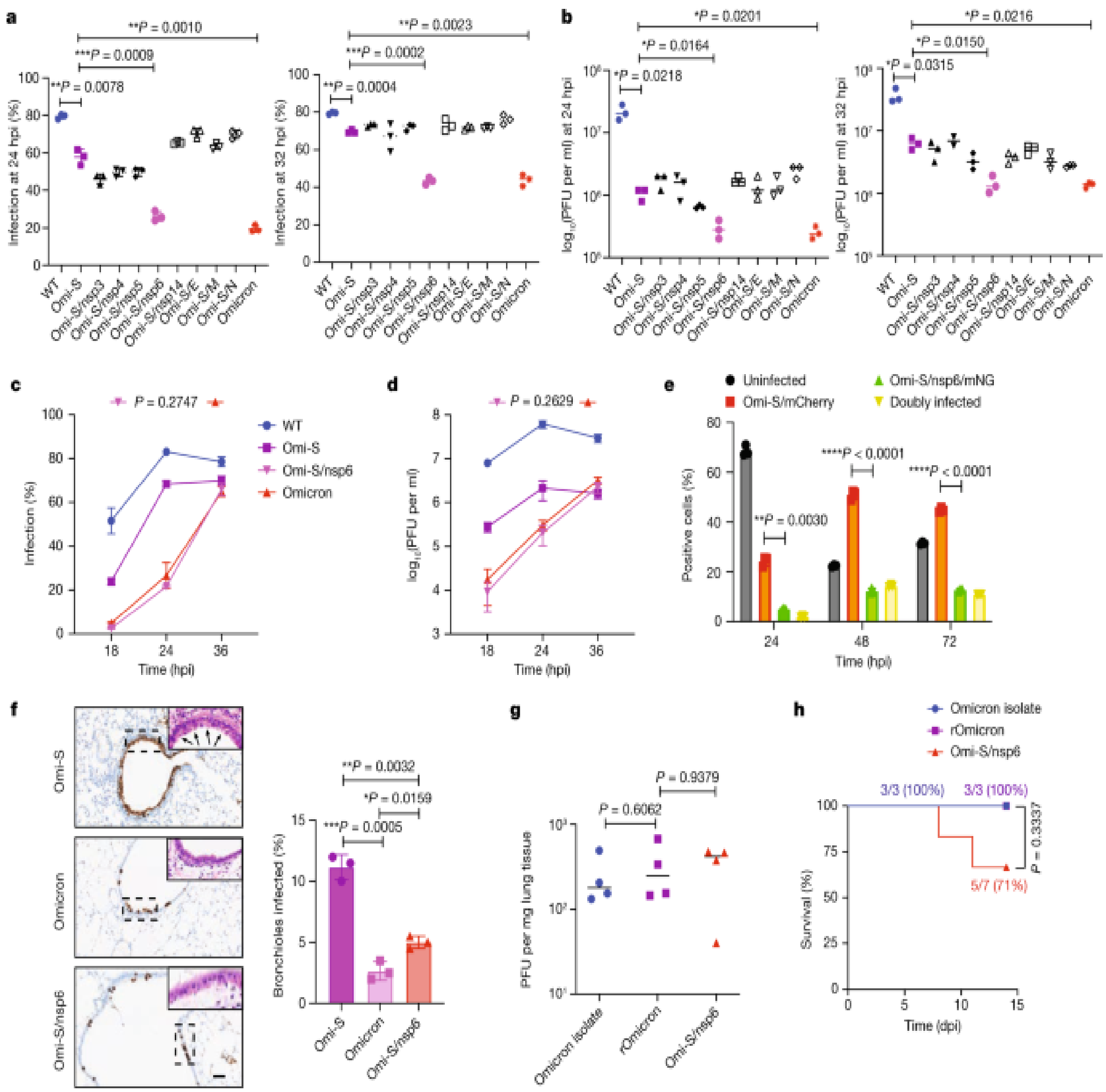
To evaluate the viral pathogenicity in lungs and nasal turbinates of K18-hACE2 mice, we performed histopathological analysis of these tissues at 2 dpi. As previously reported[4,31], an extensive near-diffused immunoreactivity of the SARS-CoV-2 N protein was detected in lung alveoli of mice infected with WT virus (Fig. 2e). By contrast, Omi-S and Omicron infection produced localized foci of alveolar staining with fewer foci for Omicron compared with Omi-S. The most marked phenotype was seen in bronchiolar epithelium, in which Omi-S caused pronounced, routinely circumferential infection, with around 10–15% of bronchioles being positive for viral N protein at 2 dpi, whereas only 3–5% of bronchioles were N-positive for Omicron (Fig. 2f). WT virus infected around 1% of bronchioles and in all

cases only included a single isolated epithelial cell per bronchiole. Furthermore, bronchiolar infection was associated with epithelial necrosis in Omi-S-infected mice, as determined through serial haematoxylin and eosin (H&E) section analysis, whereas no histological evidence of airway injury was observed in Omicron- or WT-infected mice (Extended Data Fig. 8a,b). The nasal turbinates of mice inoculated with WT and Omi-S viruses both contained abundant SARS-CoV-2-positive cells, which were associated with overt cytopathic effects, whereas Omicron produced rare, sporadic positive cells, with no apparent signs of epithelial injury (Extended Data Fig. 8c). Overall, these findings suggest that replication of Omicron in the mice respiratory tract is substantially attenuated compared to Omi-S, supporting our conclusion that mutations in S are only partially responsible for the attenuated pathogenicity of Omicron.

Mutations in S and nsp6 define Omicron attenuation

In addition to the S protein, Omicron has amino acid changes in non-structural protein 3 (nsp3), nsp4, nsp5, nsp6, nsp14, envelope (E), membrane (M) and N proteins, when compared with WT virus (Extended Data Fig. 9a). To identify non-spike proteins that are involved in Omicron attenuation, we generated a large panel of fluorescently labelled chimeric viruses, each containing Omicron S in combination with one non-spike protein of Omicron, with the remaining proteins coming from WT virus (Extended Data Fig. 9b). When we combined Omicron S with Omicron nsp6 (Omi-S/nsp6), we observed a strong decrease in viral replication, with infection kinetics mimicking those of Omicron in cell culture (Fig. 3a–d); no such decrease was seen for other chimeric viruses. Poor replication efficiency of Omi-S/nsp6 was also corroborated by our finding that both Omi-S/nsp6 and Omicron took almost five to six days to recover by CPER, whereas all other variants were recovered in two days (data not shown). Finally, like Omicron, Omi-S/nsp6 was clearly outcompeted by Omi-S in a direct competition assay (Fig. 3e).

Fig. 3: Mutations in S and nsp6 drive Omicron pathogenicity.



a–**d**, Replication kinetics of indicated mNeonGreen reporter viruses in ACE2/TMPRSS2/Caco-2 cells (MOI = 0.01) determined by flow cytometry ($n = 3$ replicates) (**a**,**c**) and plaque assay ($n = 3$ replicates) (**b**,**d**). Experiments were repeated twice. **e**, ACE2/TMPRSS2/Caco-2 cells were infected with virus mixtures at a 1:1 ratio to obtain the final MOI of 0.005 for each virus. The cells were fixed at the indicated times and analysed by flow cytometry. Percentage of uninfected, singly infected and doubly infected cells is shown. Singly infected cells were used for compensation. Individual data points are plotted along with the mean ± s.d. ($n = 3$ replicates). The experiment was repeated twice. **f**–**h**, K18-hACE2 mice were intranasally inoculated with 1×10^4 PFU of viruses. Lung samples of infected mice were collected at 2 dpi

for IHC detection of the N protein ($n = 3$ mice) (**f**) or for determination of viral titres ($n = 4$ mice) (**g**). In **f**, representative images of H&E staining of N-positive bronchioles are shown in insets. Bronchiolar epithelial necrosis is indicated with arrows. No evidence of necrosis was seen in bronchioles of mice infected with Omicron or Omi-S/nsp6. Scale bar, 100 µm. The right graph in **f** shows the percentage of N-positive bronchioles in the lungs of infected mice. Each dot represents an infected mouse. **h**, Survival of infected mice monitored daily for 14 days. Mice that lost 20% of their initial body weight were euthanized. Statistical significance was determined using a two-tailed, unpaired *t*-test with Welch's correction (**a**–**g**) and log-rank (Mantel-Cox) test (**h**). *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ and ****$P < 0.0001$; NS, not significant.

Source data

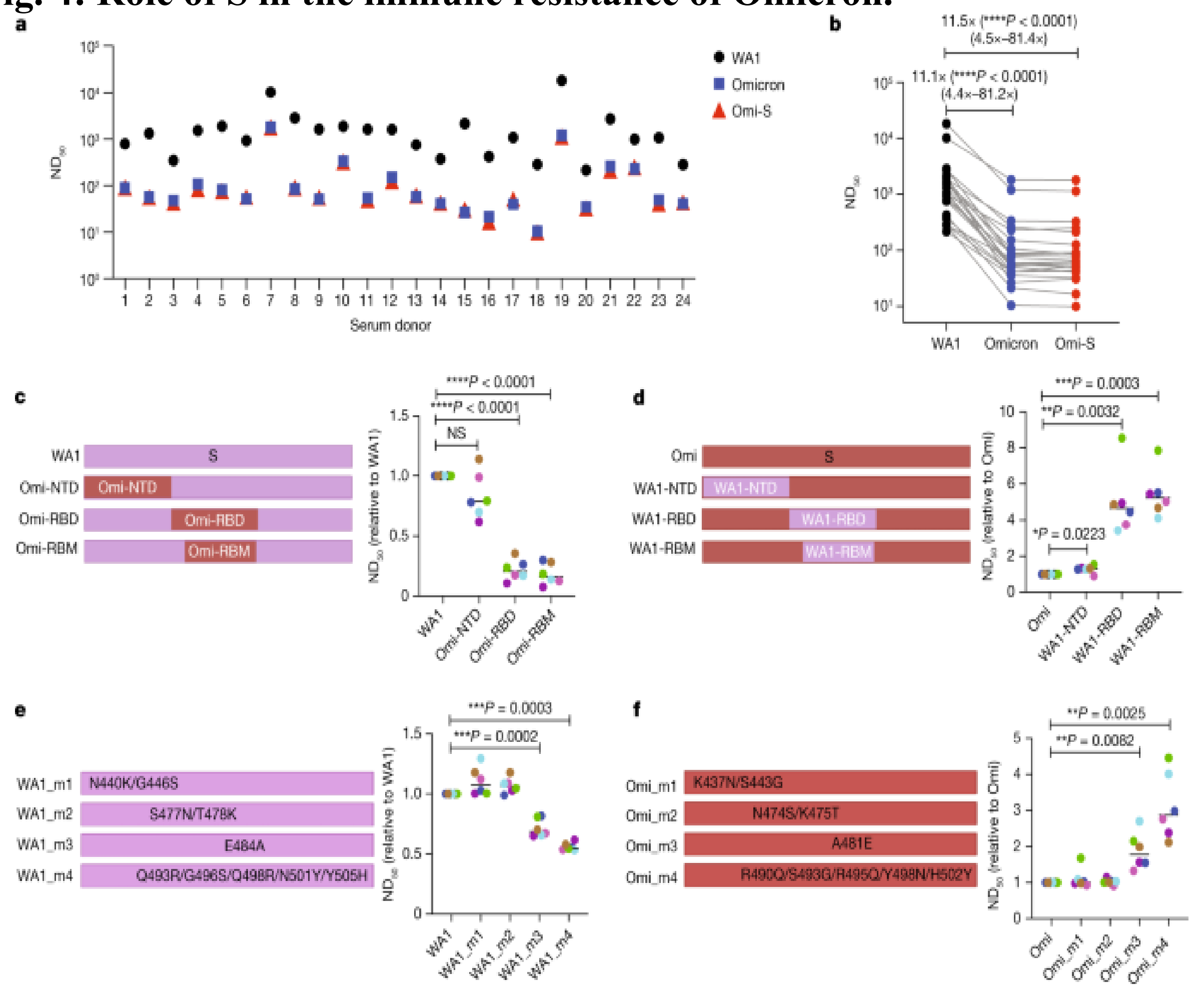
In lungs of K18-hACE2 mice, whereas Omi-S caused extensive bronchiolar infection and injury, both Omicron and Omi-S/nsp6 showed decreased infection with no evidence of epithelial damage (Fig. 3f). Consistent with these findings, lungs of Omi-S/nsp6-infected mice produced viral titres equivalent to those seen for rOmicron and Omicron isolate (Fig. 3g). Finally, 71% of mice infected with Omi-S/nsp6 survived (Fig. 3h)—in contrast with the survival rates of only 20% that were observed in mice infected with Omi-S (Fig. 2c). Overall, these results indicate that mutations in S and nsp6 are sufficient to define Omicron's attenuated virulence. These observations support and further extend the findings of a previous study showing that mutations in the 5′-UTR–nsp12 region, in which nsp6 resides, contribute to Omicron's attenuation in K18-hACE2 mice[30].

The S protein RBM drives Omicron's vaccine escape

A large body of literature has provided evidence of the extensive escape of Omicron from vaccine-induced humoral immunity[14,19]. To define S regions that are associated with the immune-escape phenotype of Omicron, we first compared the in vitro neutralization activity of sera from vaccinated individuals against WT SARS-CoV-2 (USA-WA1/2020), Omi-S and Omicron. Sera collected within two months of the second dose of mRNA-

1273 (Moderna mRNA vaccine; $n = 12$) or BNT162b2 (Pfizer-BioNTech mRNA vaccine; $n = 12$) vaccine were included (Extended Data Table 1). We performed a multicycle neutralization assay using a setting in which the virus and neutralizing sera were present at all times, mimicking the situation in a seropositive individual. All sera poorly neutralized Omicron, with an 11.1-fold (range: 4.4-fold to 81.2-fold; $P < 0.0001$) lower half-maximal neutralizing dilution (ND_{50}) for Omicron compared with WA1 (Fig. 4a,b). In fact, around 80% of samples did not completely neutralize Omicron at the highest tested concentration (Extended Data Fig. 10). Notably, Omi-S exhibited identical ND_{50} values to Omicron (11.5-fold lower than that of WA1; $P < 0.0001$) (Fig. 4a,b), suggesting that the Omicron S protein, when incorporated into a WT virus, behaves the same way as it does in Omicron.

Fig. 4: Role of S in the immune resistance of Omicron.



a, ND_{50} values for WA1, Omi-S and Omicron in sera from individuals who received two shots of Moderna (donors 1–12) or Pfizer (donors 13–24) vaccine (further details of sera are provided in Extended Data Table 1; individual curves are shown in Extended Data Fig. 10). **b**, Trajectories of ND_{50} values against WA1, Omi-S and Omicron (the data from **a** are plotted). The fold change in ND_{50} values is indicated ($n = 24$ serum samples). **c**–**f**, Schematic of the chimeric (left; **c**,**d**) and mutant (left; **e**,**f**) viruses. The amino acid numbering for WA1 mutants in **e** is based on the WA1 S sequence, whereas the numbering for Omicron mutants in **f** is based on the Omicron S sequence. Six of the 24 sera (3 from Moderna and 3 from Pfizer) were tested. Each serum sample is represented by a dot of specific colour. The data are plotted as the fold change of the parental virus. Statistical significance was determined using a two-tailed, unpaired *t*-test with Welch's correction. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ and ****$P < 0.0001$; NS, not significant.

The SARS-CoV-2 S protein comprises two domains: the S1 domain, which interacts with the ACE2 receptor, and the S2 domain, which is responsible for membrane fusion[32]. Within the S1 domain lie an N-terminal domain (NTD) and a receptor-binding domain (RBD), which contains the receptor-binding motif (RBM) that makes direct contact with the ACE2 receptor[33]. The NTD of Omicron S has 11 amino acid changes, including 6 deletions and one 3-amino-acid-long insertion, whereas the RBD contains 15 mutations, 10 of which are concentrated in the RBM (Extended Data Fig. 1). Both the NTD and the RBD host neutralizing epitopes[34,35,36,37], but the RBD is immunodominant and represents the primary target of the neutralizing activity that is present in SARS-CoV-2 immune sera[37,38]. To determine whether the neutralization resistance phenotype of Omicron is caused by mutations in a particular domain of the S protein, we generated two groups of chimeric viruses. The first group comprised the WA1 virus carrying the NTD, RBD or RBM of Omicron (Fig. 4c), and the second group consisted of Omi-S virus bearing the NTD, RBD or RBM of WA1 (Fig. 4d). The neutralization assay showed that mutations in the RBM were the main cause of Omicron's resistance to vaccine-induced humoral immunity: replacing the RBM of WA1 with that of Omicron decreased the ND_{50} by 5.4-fold ($P < 0.0001$), and, conversely, substituting the RBM of Omi-S with

that of WA1 increased the ND_{50} by 5.6-fold (*P* = 0.0003) (Fig. 4c,d). The fact that none of the RBM-swap viruses achieved the difference of around 11-fold that was seen between WA1 and Omi-S suggests that mutations in other parts of S also contribute to vaccine resistance.

To investigate whether specific mutations in Omicron RBM drive vaccine escape, we generated two additional panels of recombinant viruses, one with WA1 S carrying Omicron RBM mutations, either singly or in combination (Fig. 4e), and the other with Omicron S lacking the same set of mutations (Fig. 4f). Two WA1 mutants—mutant 3 (with an E484A substitution) and mutant 4 (bearing a cluster of five substitutions Q493R, G496S, Q498R, N501Y and Y505H)—exhibited a moderate but statistically significant decrease of 1.4-fold (*P* = 0.0002) and 1.8-fold (*P* = 0.0003) in ND_{50} values, respectively, compared with WA1 (Fig. 4e). The opposite was observed when these mutations were removed from Omicron S; the Omicron mutant 3 (lacking the E484A substitution) and mutant 4 (lacking Q493R, G496S, Q498R, N501Y and Y505H) had 1.9-fold (*P* = 0.0082) and 3.1-fold (*P* = 0.0025) higher ND_{50} values compared with Omicron (Fig. 4f). As none of the mutants captured the overall phenotype of Omicron, we assume that the vaccine escape is a cumulative effect of mutations distributed along the length of the S protein. It is possible that mutations alter the conformation of Omicron S in such a manner that most of the immunodominant neutralizing epitopes are disrupted and become unavailable for neutralization.

Discussion

This study provides key insights into viral proteins that contribute to SARS-CoV-2 pathogenicity. We show that S, the most mutated protein in Omicron, has an incomplete role in Omicron attenuation. In cell-based infection assays, the Omi-S virus exhibits an intermediate replication efficiency between the ancestral virus and Omicron. Similarly, in K18-hACE2 mice, Omi-S contrasts with non-fatal Omicron and leads to 80% mortality; the ancestral virus causes 100% mortality in these mice. Notably, when we combined S mutations with mutations in nsp6, the virus exhibited an attenuated phenotype largely resembling that of Omicron, indicating that these two proteins are major determinants of Omicron pathogenicity. Future

studies will decipher the mechanism(s) by which nsp6 mutations affect viral replication.

One potential limitation of our study is the use of K18-hACE2 mice for pathogenesis studies instead of primate models that have more similarities with humans[39]. It should however be noted that K18-hACE2 mice are a well-established model for investigating the lethal phenotype of SARS-CoV-2[4,31]. Although these mice develop lung pathology after SARS-CoV-2 infection, mortality has been associated with involvement of the central nervous system owing to viral neuroinvasion and dissemination[31,40]. The fact that infection of K18-hACE2 mice with Omi-S, but not with Omicron, elicits neurological signs (for example, hunched posture and a lack of responsiveness) suggests that the neuroinvasion property is preserved in Omi-S, probably as a result of its higher replication efficiency, and that the determinants of this property lie outside of the S protein. These findings are consistent with a previous study showing that hamsters infected with Omi-S shed significantly more virus and lost more weight than those infected with Omicron, suggesting that mutations outside of S contribute to the attenuated pathogenicity of Omicron[41].

We found that although the ancestral virus mainly replicates in lung alveoli and causes only rare infection of bronchioles in K18-hACE2 mice, both Omi-S and Omicron exhibit an increased propensity to replicate in bronchiolar epithelium, indicating that the S protein is accountable for the changed tropism. The mechanism behind this switch is unknown, but it is possible that Omicron S is more efficient than WT S in using cathepsin B or cathepsin L (refs. [25,42,43]), which form an active viral entry pathway in bronchioles and other airway cells[41]. By contrast, the entry of SARS-CoV-2 into alveolar epithelial cells is mainly driven by TMPRSS2 (refs. [28,44]), which Omicron S is deficient in utilizing[25,45], leading to poor infection of these cells[4,5,25,43]. These findings may explain the attenuated lung pathology caused by Omicron.

Omicron nsp6 has two altered sites relative to the prototype SARS-CoV-2 Wuhan-Hu-1 isolate: a three-amino-acid deletion (LSG, positions 105–107) and an I189V substitution (Extended Data Fig. 9). Several functions of nsp6 in coronavirus replication have been described; chief among them is the

biogenesis of double-membrane vesicles (DMVs), which represent the site of viral RNA synthesis[46,47,48,49,50]. A previous study showed that SARS-CoV-2 DMVs are mainly generated by the concerted action of three viral proteins: nsp3, nsp4, and nsp6; although nsp3 and nsp4 are sufficient for the formation of DMVs, nsp6 connects these DMVs with the endoplasmic reticulum and channelizes the essential communication between these structures[46]. Whether the constellation of mutations in Omicron nsp6 affects the formation or functions of DMVs needs further investigation. Nsp6 also activates NLR3-dependent cytokine production and pyroptosis in the lungs of patients with COVID-19, serving as a key virulence factor[47]. Of note, an nsp6 variant that is associated with asymptomatic COVID-19 exhibited a reduced ability to induce pyroptosis[47], prompting speculation that mutations in Omicron nsp6 may also influence pyroptosis. Detailed mechanistic studies will be required to dissect the effect of Omicron mutations on the functions of nsp6.

It is at present unknown whether mutations in S and nsp6 work in concert with each other to drive Omicron attenuation. Given that Omicron S showed a higher predilection for bronchioles, it is possible that S is responsible for the altered viral tropism, whereas non-spike mutations—including those in nsp6—are mere adaptations to the changed tissue environment. It is worth mentioning that although nsp6 seems to be the major non-spike protein behind Omicron attenuation, the contribution of other viral proteins cannot be completely ruled out. In vitro experiments examining the role of non-spike mutations were all carried out in ACE2/TMPRSS2/Caco-2 cells. Using other, more immune-competent cell types could reveal the effect of other non-spike mutations as well. In addition, our chimeric viruses contained Omicron S paired with only one non-spike protein at a time, which limited long-range epistatic interactions between mutations in multiple viral proteins.

Our study shows that mutations in the RBM of Omicron S are the main determinants of Omicron's escape from neutralizing antibodies, although mutations in other regions of S also contribute. Within the RBM, we identify two hotspots of mutations, which give Omicron S the ability to resist neutralization: one bearing the E484A substitution and the other containing a cluster of five substitutions—Q493R, G496S, Q498R, N501Y and Y505H.

The E484A substitution has been shown to escape neutralization by convalescent sera[51]. Moreover, structural modelling suggests that some therapeutic monoclonal antibodies establish highly stable salt bridges with the E484 residue, entirely losing their binding when this residue is changed to A or after Q493K and Y505H changes[52]. Similarly, mapping of RBM residues that directly interact with 49 known neutralizing antibodies revealed N440, G446, S477 and T478 as low-frequency interactors, N501, Y505 and Q498 as medium-frequency interactors and E484 and Q493 as high-frequency interactors[53], in line with our neutralization assay results. Notably, although the antibody-binding potential of Omicron S is impaired[54], its receptor-binding capacity is intact. In fact, the Omicron RBD has a higher affinity for ACE2 relative to the Wuhan-Hu-1 and Delta RBDs[25]. This indicates that mutations in Omicron S have evolved in such a manner that they hinder antibody binding but preserve the receptor engagement. This opens up the possibility of targeting the conserved and structurally constrained regions of S that are involved in ACE2 recognition for the design of broad-spectrum vaccines to control the COVID-19 pandemic.

Methods

Cells, antibodies and plasmids

The cell lines were incubated at 37 °C and 5% CO_2 in a humidified incubator. Human embryonic kidney HEK293T cells (ATCC; CRL-3216), human lung adenocarcinoma A549 cells (ATCC; CCL-185), human colorectal adenocarcinoma Caco-2 cells (ATCC; HTB-37) and African green monkey kidney Vero E6 cells were maintained in Dulbecco’s modified Eagle’s medium (DMEM) (Gibco; 11995-065) containing 10% fetal bovine serum (FBS) and 1× non-essential amino acids. A lentiviral delivery system was used to generate cells stably expressing human ACE2 and TMPRSS2. The mycoplasma-negative status of all cell lines was confirmed.

Anti-SARS-CoV N protein antibody (Rockland; 200-401-A50; 1:2,000) was used to detect the SARS-CoV-2 N protein by immunofluorescence and western blot. Mouse anti-SARS-CoV-2 spike antibody (GeneTex;

GTX632604; 1:1,000), directed against the S2 subunit, was used for western blot analysis of S-protein cleavage in virus particles and infected cells.

Plasmids encoding various fragments of the SARS-CoV-2 genome (Hu/DP/Kng/19-020 isolate) were a gift from Y. Matsuura[55]. We replaced the S gene in the plasmid pCSII-SARS-CoV-2 F8 (ref. [55]) with the chemically synthesized Omicron S gene and named this plasmid pCSII-SARS-CoV-2 F8_Omicron. We replaced the open reading frame (ORF) 7 in the plasmid pcDNA3.1-SARS-CoV-2 F9+10 with mNeonGreen or mCherry to obtain plasmids pcDNA3.1-SARS-CoV-2 F9+10_mNG and pcDNA3.1-SARS-CoV-2 F9+10_mCherry, respectively. The plasmids pMW-CoV-2-UTRlinker (ref. [55]) and pGL-CPERlinker (ref. [24]), both containing a linker fragment comprising hepatitis delta virus ribozyme (HDVr), the bovine growth hormone polyadenylation signal sequence (BGH-polyA) and cytomegalovirus (CMV) promoter, have been previously reported. The lentiviral vectors, pLOC_hACE2_PuroR and pLOC_hTMPRSS2_BlastR, containing human ACE2 and TMPRSS2, respectively, have been described previously[56].

Biocontainment

All procedures were performed in a state-of-the-art biosafety level 3 (BSL3) facility at the National Emerging Infectious Diseases Laboratories (NEIDL) of Boston University using biosafety protocols approved by the Institutional Biosafety Committee (IBC). The experimental plans, including the generation of recombinant chimeric viruses, were reviewed and approved by the IBC, which comprises scientists, biosafety and compliance experts and members of the local community. Furthermore, the research was approved by the Boston Public Health Commission. All personnel received rigorous biosafety, biosecurity and BSL3 training before participating in experiments. Personal protective equipment, including scrubs, disposable overalls, shoe covers, double-layered gloves and powered air-purifying respirators, was used. Biosecurity measures are built in the environment through building and security systems and are reinforced through required training programmes, standing meetings and emergency exercises. The researchers involved in working with chimeric viruses received at least two booster shots of the SARS-CoV-2 mRNA vaccine before the study was started.

Finally, all researchers were medically cleared by the Boston University Research Occupational Health Program.

Collection of serum samples

Sera from individuals who received two doses of mRNA-1273 (Moderna) or BNT162b2 (Pfizer) vaccine were collected at Boston Medical Center at least two weeks after the final dose. These individuals had no prior history of SARS-CoV-2 infection. Serum samples were collected using protocols reviewed and approved by the Institutional Review Board at Boston Medical Center. All methods were performed in accordance with relevant guidelines and regulations. The participants provided electronic informed consent. De-identified samples were used in this research. Additional information for serum samples is provided in Extended Data Table 1.

Omicron stock preparation and titration

The SARS-CoV-2 BA.1 Omicron virus stock was generated in ACE2/TMPRSS2/Caco-2 cells. In brief, 5×10^5 cells, grown overnight in DMEM, 10% FBS and 1× NEAA in one well of a six-well plate, were inoculated with the collection medium in which the nasal swab from a patient with COVID-19 was immersed. The swab material was obtained from the Department of Public Health, Massachusetts, and it contained the sequence-verified Omicron virus (NCBI accession number: OL719310). Twenty-four hours after infecting cells, the culture medium was replaced with 2 ml of DMEM, 2% FBS and 1× NEAA and the cells were incubated for another 72 h, at which point the CPE became visible. The culture medium was collected, passed through a 0.45 μm filter and kept at −80 °C as a P0 virus stock. To generate a P1 stock, we infected 1×10^7 ACE2/TMPRSS2/Caco-2 cells, seeded the day before in a T175 flask, with the P0 virus at an MOI of 0.01. The next day, the culture medium was changed to 25 ml of 2% FBS-containing medium. Three days later, when the cells exhibited excessive CPE, the culture medium was collected, passed through a 0.45 μm filter and stored at −80 °C as a P1 stock.

To titrate the virus stock, we seeded ACE2/TMPRSS2/Caco-2 cells into a 12-well plate at a density of 2×10^5 cells per well. The next day, the cells

were incubated with serial 10-fold dilutions of the virus stock (250 µl volume per well) for 1 h at 37 °C, overlaid with 1 ml per well of medium containing a 1:1 mixture of 2× DMEM/4% FBS and 1.2% Avicel (DuPont; RC-581) and incubated at 37 °C for another three days. To visualize the plaques, the cell monolayer was fixed with 4% paraformaldehyde and stained with 0.1% crystal violet, with both fixation and staining performed at room temperature for 30 min each. The number of plaques was counted and the virus titre was calculated.

Generation of recombinant SARS-CoV-2 by CPER

SARS-CoV-2 recombinant viruses were generated using a previously described optimized CPER protocol[24]. Full-length SARS-CoV-2 cDNA cloned into a bacterial artificial chromosome (BAC)[23] was used to generate WT and Omi-S viruses. In brief, the BAC was amplified into eight overlapping fragments (F1, F2, F3, F4, F5, F6, F7 and F9) covering the whole SARS-CoV-2 genome. The pCSII-SARS-CoV-2 F8 (containing a D614G substitution) and pCSII-SARS-CoV-2 F8_Omicron plasmids, which were used to generate S mutants, served as templates for the amplification of fragment 8 (F8). The UTR linker plasmids pMW-CoV-2-UTRlinker (ref. [55]) or pGL-CPERlinker (ref. [24]) were used as a template to amplify the linker sequence. The 5′ termini of all ten DNA fragments (F1–F9 and the linker) were phosphorylated by using T4 PNK (NEB; M0201). The CPER reaction containing equimolar amounts (0.05 pmol) of each fragment was performed with PrimeStar GXL DNA polymerase (Takara Bio; R050A) as previously described[24]. The nicks in the circular product were sealed by using HiFi Taq DNA ligase (NEB; M0647S).

To generate chimeric viruses containing a combination of Omicron S and non-spike proteins (Omi-S/nsp3, Omi-S/nsp4, Omi-S/nsp5, Omi-S/nsp6, Omi-S/nsp14, Omi-S/E, Omi-S/M and Omi-S/N), we used SARS-CoV-2 plasmids described previously[55] as templates, provided by Y. Matsuura. These plasmids contained SARS-CoV-2 sequences derived from the SARS-CoV-2/Hu/DP/Kng/19-020 strain. We introduced mutations into these plasmids using the standard DNA recombination technology. Our chimeric viruses also contained a P323L substitution in nsp12. Plasmid sequences were confirmed by the Sanger method.

To transfect cells with the CPER product, we seeded ACE2/TMPRSS2/Caco-2 cells into a six-well plate at a density of 5×10^5 cells per well. The transfection mix was prepared by mixing 26 µl of the original 52-µl CPER reaction volume with 250 µl of Opti-MEM (Thermo Fisher Scientific; 31985070) and 6 µl of TransIT-X2 Dynamic Delivery System (Mirus Bio; MIR 6000). After incubation at room temperature for 25 min, the transfection mix was added to the cells. The next day, the culture medium was replaced with fresh DMEM containing 2% FBS. The CPE became visible in three to four days, at which point the culture medium was collected and stored as a P0 virus stock. The P0 stock was used for experiments described in this manuscript. The sequence of CPER-generated viruses was confirmed by next-generation sequencing.

SARS-CoV-2 whole-viral sequencing and genome assembly

cDNA synthesis was performed using Superscript IV reverse transcriptase (Invitrogen). Whole-viral amplification was performed with the NEB Varskip protocol using multiplexed primer pools designed with Primal Scheme, generating 400-bp tiling amplicons. PCR products from the Varskip protocol were pooled together and Illumina library construction was performed using the Nextera XT Library Prep Kit (Illumina). Deep-sequencing data analysis was performed using the Stanford Coronavirus Antiviral & Resistance Database (CoV-RDB) platform[57]. Input FASTQ sequence alignment with the Wuhan-Hu-1 reference was done using MiniMap2 v.2.22 in the CodFreq pipeline (https://github.com/hivdb/codfreq). The output of MiniMap2, an aligned SAM file, was converted to a CodFreq file by a Python script written in-house, using a PySam library (v.0.18.0), and was further analysed with the CoV-RDB. SARS-COV-2 variant calling was done using CoV-RDB Scorpio call v.1.2.123 (ref. [58]) (https://pangolin.cog-uk.io/) and Nextclade v.1.13.2 (ref. [59]) (https://clades.nextstrain.org/). PCR and sequencing run were performed once with the appropriate positive and negative controls.

SARS-CoV-2 neutralization assay

For neutralization assays, initial 1:10 dilutions of plasma, obtained from individuals who received two shots of either the Moderna or the Pfizer mRNA-based SARS-CoV-2 vaccine, were fivefold serially diluted in Opti-MEM over seven or eight dilutions. These plasma dilutions were then mixed at a 1:1 ratio with 1.25×10^4 infectious units of SARS-CoV-2 and incubated for 1 h at 37 °C. Thereafter, 100 µl of this mixture was directly applied to ACE2/A549 cells seeded the previous day in poly-l-lysine-coated 96-well plates at a density of 2.5×10^4 cells per well in a volume of 100 µl. Thus, the final starting dilution of plasma was 1:20 and the final MOI was 0.5. The cells were incubated at 37 °C for 24 h, after which they were fixed and stained with an anti-N antibody. When PBS instead of plasma was used as a negative control, these infection conditions resulted in around 40–50% infected cells at 24 hpi.

Generation and infection of iAT2 cells

The detailed protocol for generation of human iPS-cell-derived alveolar epithelial type II cells (iAT2 cells) has been published in our previous papers[28,60]. The ALI cultures were established by preparing single-cell suspensions of iAT2 three-dimensional (3D) sphere cultures grown in Matrigel. In brief, Matrigel droplets containing iAT2 spheres were dissolved in 2 mg ml^{-1} dispase (Sigma) and the spheres were dissociated in 0.05% trypsin (Gibco) to generate a single-cell suspension. Transwell inserts (6.5 mm; Corning) were coated with dilute Matrigel (Corning) in accordance with the manufacturer's protocol. Single-cell iAT2 cells were plated on Transwells at a density of 520,000 cells per cm^2 in 100 µl of CK+DCI medium containing 10 µM of rho-associated kinase inhibitor ("Y"; Sigma Y-27632). Six hundred microlitres of this medium was added to the basolateral compartment. Twenty-four hours after plating, the basolateral medium was changed with fresh CK+DCI+Y medium. Forty-eight hours after plating, the apical medium was aspirated to initiate the ALI culture. Seventy-two hours after plating, the basolateral medium was replaced with CK+DCI medium to remove the rho-associated kinase inhibitor. Basolateral medium was changed every two days thereafter. The detailed composition of CK+DCI medium is provided in our previous publications[28,60].

iAT2 cells in ALI cultures were infected with purified SARS-CoV-2 stock at an MOI of 2.5 based on the titration that was done on ACE2/TMPRSS2/Caco-2 cells. For infection, 100 µl of inoculum prepared in 1× PBS (or mock-infected with PBS only) was added to the apical chamber of each Transwell and incubated for 2 h at 37 °C, followed by the removal of the inoculum and washing of the apical side three times with 1× PBS (100 µl per wash). The cells were incubated for two or four days, after which the newly released virus particles on the apical side were collected by adding 100 µl of 1× PBS twice to the apical chamber and incubating at 37 °C for 15 min. The number of infectious virus particles in the apical washes was measured by the plaque assay on ACE2/TMPRSS2/Caco-2 cells. For flow cytometry, iAT2 cells were detached by adding 0.2 ml Accutase (Sigma; A6964) apically, and incubated at room temperature for 15 min. The detached cells were pelleted by low-speed centrifugation, fixed in 10% formalin and stained with anti-SARS-CoV-2 N antibody.

Detection of S-protein incorporation and cleavage in SARS-CoV-2 particles

The culture medium of ACE2/TMPRSS2/Caco-2 cells transfected with the CPER product was collected and passed through 0.22-µm filters. The SARS-CoV-2 particles were pelleted down by mixing an equal volume of the culture medium with 20% PEG6000 in PBS followed by overnight incubation at 4 °C. The mixture was centrifuged at 12,000*g* for 30 min at 4 °C, and the pellet was resuspended in 1× SDS sample buffer. The protein concentration was measured by the BCA assay using Pierce BCA Protein Assay kit (Thermo Fisher Scientific; 23225). Equal amounts of protein were resolved on 4–12% SDS page. The S protein was detected with mouse anti-SARS-CoV-2 spike antibody (GeneTex; GTX632604; 1:1,000) and IRDye 800CW donkey anti-mouse IgG secondary antibody (LI-COR Biosciences; 926-32212; 1:5,000). The bands were visualized by scanning the membrane with the LiCor CLx infrared scanner. The open-source package ImageJ (v.1.53a) was used to measure the intensity of protein bands.

Flow cytometry

For flow cytometry, fixed cells were permeabilized in 1× permeabilization buffer (Thermo Fisher Scientific; 00-5523-00) and stained with SARS-CoV-2 nucleocapsid antibody (Rockland; 200-401-A50, 1:1,000), followed by donkey anti-rabbit IgG-AF647 secondary antibody (Thermo Fisher Scientific; A-31573). Cells infected with fluorescent reporter viruses were fixed and analysed without staining. Gating was based on uninfected, stained control cells. The extent of staining was quantified using a BD LSR II flow cytometer (BD Biosciences), and the data were analysed with FlowJo v.10.6.2 (FlowJo, Tree Star). The gating strategy for flow cytometry is shown in Supplementary Fig. 1.

Immunofluorescence

Immunofluorescence was performed as described in our previous publication[56]. In brief, virus-infected cells were fixed in 4% paraformaldehyde and permeabilized in a buffer containing 0.1% Triton X-100 prepared in PBS. After blocking in a buffer containing 0.1% Triton X-100, 10% goat serum and 1% BSA, the cells were incubated overnight at 4 °C with anti-SARS-CoV N antibody (1:2,000 dilution). The cells were then stained with Alexa Fluor 568-conjugated goat anti-rabbit secondary antibody (1:1,000 dilution) (Invitrogen; A11008) in the dark at room temperature for 1 h and counterstained with DAPI. Images were captured using the ImageXpress Micro Confocal (IXM-C) High-Content Imaging system (Molecular Devices) with a 4× S Fluor objective lens at a resolution of 1.7 µm per pixel in the DAPI (excitation, 400 nm/40 nm; emission, 447 nm/60 nm) and TexasRed (excitation, 570 nm/80 nm; emission, 624 nm/40 nm) channels. Both channels were used to establish their respective laser autofocus offsets. The images were analysed using MetaXpress High Content Image Acquisition and Analysis software (Molecular Devices). First, the images were segmented using the CellScoring module. The objects between 7 µm and 20 µm in diameter and greater than 1,800 grey level units in intensity were identified and classified as nuclei. Positive cells were taken as nuclei having a TexasRed signal of 1,500 grey level units or above within 10–20 µm of each nucleus. The remaining objects were set to negative cells. From these objects, the following readouts were measured and used for downstream analysis: total number of positive and negative cells; total area of positive cells; and integrated intensity in the TexasRed channel for

positive cells. To calculate the ND_{50}, we performed a non-linear regression curve fit analysis using Prism 9 software (GraphPad).

Mice maintenance and approvals

Mice were maintained in a facility accredited by the Association for the Assessment and Accreditation of Laboratory Animal Care (AAALAC). Animal studies were performed following the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocols were approved by the Boston University Institutional Animal Care and Use Committee (IACUC). Heterozygous K18-hACE2 C57BL/6J mice (Strain 2B6.Cg-Tg(K18-ACE2)2Prlmn/J) were purchased from the Jackson Laboratory (JAX). Mice were housed in groups of four or five in ventilated cages (Tecniplast) and maintained on a 12:12 light cycle at 30–70% humidity, 20 °C temperature, ad libitum water and ad libitum standard chow diet (LabDiet).

Mice infection

Male and female K18-hACE2 mice (12–20 weeks old) were inoculated intranasally with 10^4 PFU of SARS-CoV-2 in 50 µl of sterile 1× PBS. The inoculations were performed under 1–3% isoflurane anaesthesia. In vivo experiments were not blinded, and mice were randomly assigned to infection groups. No a priori sample size calculation was performed. Instead, samples sizes were determined on the basis of our previous animal work. Twenty-six mice (6 for WT, 10 for Omi-S and 10 for Omicron) were enrolled in a 14-day survival study, and another 42 mice (14 for each of the WT, Omi-S and Omicron viruses) were used for virological and histological analysis of infected lungs. During the survival study, the mice were monitored for body weight, respiration, general appearance, responsiveness and neurological signs. An IACUC-approved clinical-scoring system was used to monitor disease progression and define humane end-points. The score of 1 was given for each of the following situations: body weight, 10–19% loss; respiration, rapid and shallow with increased effort; appearance, ruffled fur and/or hunched posture; responsiveness, low to moderate unresponsiveness; and neurological signs, tremors. The sum of these individual scores constituted the final clinical score. Mice were considered moribund and humanly

euthanized in case of weight loss greater than or equal to 20%, or if they received a clinical score of 4 or above for two consecutive days. Body weight and clinical score were recorded once per day for the duration of the study. For the purpose of survival curves, mice euthanized on a given day were counted dead the day after. Mice that were found dead in the cage were counted dead on the same day. For euthanasia, an overdose of ketamine was administered followed by a secondary method of euthanasia.

For the quantification of SARS-CoV-2 infectious particles in lungs by plaque assay, lung tissues were collected in 600 µl of RNAlater stabilization solution (Thermo Fisher Scientific; AM7021) and stored at −80 °C until analysis. Twenty to forty milligrams of tissue was placed in a tube containing 600 µl of Opti-MEM and a 5-mm stainless steel bead (Qiagen; 69989) and homogenized in the Qiagen TissueLyser II by two dissociation cycles (1,800 oscillations per min for 2 min) with a 1-min interval between cycles. The homogenate was centrifuged at 15,000*g* for 10 min at room temperature and the supernatant was transferred to a new tube. Tenfold serial dilutions of this supernatant were used for the plaque assay on ACE2/TMPRSS2/Caco-2 cells, as described above.

For IHC and histological analysis, the insufflated whole lung tissues were inactivated in 10% neutral buffered formalin at a 20:1 fixative-to-tissue ratio for a minimum of 72 h before removal from BSL3 in accordance with an approved IBC protocol. Tissues were subsequently processed and embedded in paraffin, and 5-µm sections were stained with H&E following standard histological procedures. IHC was performed using a Ventana BenchMark Discovery Ultra autostainer (Roche Diagnostics). An anti-SARS-CoV-2 S antibody (Cell Signaling Technologies; clone E5S3V) or anti-SARS-CoV-2 N antibody (Cell Signaling Technologies; clone 1C7C7) that showed equivalent immunoreactivity against WT and Omicron proteins was used to identify virus-infected cells. For the SARS-CoV-2 N antibody, given its mouse origin, an additional rabbit anti-mouse anti-Ig1 + Ig2a + IgG3 antibody (Abcam; 133469) was used to prevent non-specific binding. A HRP-conjugated goat anti-rabbit IgG polymer detection was then used to detect the viral specific antibodies (Vector Laboratories; MP-7451) and finally developed using 3,3′-diaminobenzidine (DAB) chromogen and counterstained with haematoxylin. Negative and positive controls for IHC

included blocks of uninfected and SARS-CoV-2-infected Vero E6 cells, respectively.

For the quantification of N protein in the nasal turbinate epithelium, digitalized whole slide scans were analysed using the image analysis software HALO (Indica Labs). The respiratory epithelium was manually annotated to create a layer for downstream analysis. Area quantification was performed to determine the percentages of SARS-CoV-2 N in the annotated layer, which generated the percentage of immunoreactivity output.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data supporting the conclusions of this study are reported in the paper. The raw data are available from the corresponding author with no restrictions upon reasonable request. Source data are provided with this paper.

Code availability

No code was used for data acquisition in this study.

References

1. Liu, L. et al. Striking antibody evasion manifested by the Omicron variant of SARS-CoV-2. *Nature* **602**, 676–681 (2022).

2. Planas, D. et al. Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. *Nature* **602**, 671–675 (2022).

3. Schmidt, F. et al. Plasma neutralization of the SARS-CoV-2 Omicron variant. *N. Engl. J. Med.* **386**, 599–601 (2022).

4. Shuai, H. et al. Attenuated replication and pathogenicity of SARS-CoV-2 B.1.1.529 Omicron. *Nature* **603**, 693–699 (2022).

5. Halfmann, P. J. et al. SARS-CoV-2 Omicron virus causes attenuated disease in mice and hamsters. *Nature* **603**, 687–692 (2022).

6. Lewnard, J. A. et al. Clinical outcomes associated with SARS-CoV-2 Omicron (B.1.1.529) variant and BA.1/BA.1.1 or BA.2 subvariant infection in Southern California. *Nat. Med.* **28**, 1933–1943 (2022).

7. Wolter, N. et al. Early assessment of the clinical severity of the SARS-CoV-2 omicron variant in South Africa: a data linkage study. *Lancet* **399**, 437–446 (2022).

8. Ulloa, A. C., Buchan, S. A., Daneman, N. & Brown, K. A. Estimates of SARS-CoV-2 Omicron variant severity in Ontario, Canada. *J. Am. Med. Assoc.* **327**, 1286–1288 (2022).

9. Uraki, R. et al. Characterization of SARS-CoV-2 Omicron BA.4 and BA.5 isolates in rodents. *Nature* **612**, 540–545 (2022).

10. Uraki, R. et al. Characterization and antiviral susceptibility of SARS-CoV-2 Omicron BA.2. *Nature* **607**, 119–127 (2022).

11. Dyer, O. Covid-19: Omicron is causing more infections but fewer hospital admissions than delta, South African data show. *Br. Med. J.* **375**, n3104 (2021).

12. Sigal, A. Milder disease with Omicron: is it the virus or the pre-existing immunity? *Nat. Rev. Immunol.* **22**, 69–71 (2022).

13. WHO. Tracking SARS-CoV-2 variants https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/ (2022).

14. Cele, S. et al. Omicron extensively but incompletely escapes Pfizer BNT162b2 neutralization. *Nature* **602**, 654–656 (2022).

15. Gu, H. et al. Probable transmission of SARS-CoV-2 omicron variant in quarantine hotel, Hong Kong, China, November 2021. *Emerg. Infect. Dis.* **28**, 460–462 (2022).

16. Iuliano, A. D. et al. Trends in disease severity and health care utilization during the early Omicron variant period compared with previous SARS-CoV-2 high transmission periods—United States, December 2020–January 2022. *Morb. Mortal. Wkly Rep.* **71**, 146–152 (2022).

17. CDC. COVID Data Tracker https://covid.cdc.gov/covid-data-tracker/#variant-proportions (2022).

18. Taylor, L. Covid-19: Omicron drives weekly record high in global infections. *Br. Med. J.* **376**, o66 (2022).

19. Dejnirattisai, W. et al. SARS-CoV-2 Omicron-B.1.1.529 leads to widespread escape from neutralizing antibody responses. *Cell* **185**, 467–484 (2022).

20. Suzuki, R. et al. Attenuated fusogenicity and pathogenicity of SARS-CoV-2 Omicron variant. *Nature* **603**, 700–705 (2022).

21. Smyth, D. S. et al. Tracking cryptic SARS-CoV-2 lineages detected in NYC wastewater. *Nat. Commun.* **13**, 635 (2022).

22. Kirby, A. E. et al. Notes from the field: early evidence of the SARS-CoV-2 B.1.1.529 (Omicron) variant in community wastewater—United States, November–December 2021. *Morb. Mortal. Wkly Rep.* **71**, 103–105 (2022).

23. Herrmann, A. et al. Cloning of a passage-free SARS-CoV-2 genome and mutagenesis using red recombination. *Int. J. Mol. Sci.* **22**, 10188 (2021).

24. Liu, G. & Gack, M. U. An optimized circular polymerase extension reaction-based method for functional analysis of SARS-CoV-2. Preprint at https://doi.org/10.1101/2022.11.26.518005 (2022).

25. Meng, B. et al. Altered TMPRSS2 usage by SARS-CoV-2 Omicron impacts tropism and fusogenicity. *Nature* **603**, 706–714 (2022).

26. Willett, B. J. et al. SARS-CoV-2 Omicron is an immune escape variant with an altered cell entry pathway. *Nat. Microbiol.* **7**, 1161–1179 (2022).

27. Martines, R. B. et al. Pathology and pathogenesis of SARS-CoV-2 associated with fatal coronavirus disease, United States. *Emerg. Infect. Dis.* **26**, 2005–2015 (2020).

28. Huang, J. et al. SARS-CoV-2 infection of pluripotent stem cell-derived human lung alveolar type 2 cells elicits a rapid epithelial-intrinsic inflammatory response. *Cell Stem Cell* **27**, 962–973 (2020).

29. Mulay, A. et al. SARS-CoV-2 infection of primary human lung epithelium for COVID-19 modeling and drug discovery. *Cell Rep.* **35**, 109055 (2021).

30. Liu, S., Selvaraj, P., Sangare, K., Luan, B. & Wang, T. T. Spike protein-independent attenuation of SARS-CoV-2 Omicron variant in laboratory mice. *Cell Rep.* **40**, 111359 (2022).

31. Carossino, M. et al. Fatal neurodissemination and SARS-CoV-2 tropism in K18-hACE2 mice is only partially dependent on hACE2 expression. *Viruses* **14**, 535 (2022).

32. Huang, Y., Yang, C., Xu, X. F., Xu, W. & Liu, S. W. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. *Acta Pharmacol. Sin.* **41**, 1141–1149 (2020).

33. Lan, J. et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215–220 (2020).

34. Chi, X. et al. A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science* **369**, 650–655 (2020).

35. Voss, W. N. et al. Prevalent, protective, and convergent IgG recognition of SARS-CoV-2 non-RBD spike epitopes. *Science* **372**, 1108–1112 (2021).

36. Premkumar, L. et al. The receptor binding domain of the viral spike protein is an immunodominant and highly specific target of antibodies in SARS-CoV-2 patients. *Sci. Immunol.* **5**, eabc8413 (2020).

37. Ju, B. et al. Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* **584**, 115–119 (2020).

38. Piccoli, L. et al. Mapping neutralizing and immunodominant sites on the SARS-CoV-2 spike receptor-binding domain by structure-guided high-resolution serology. *Cell* **183**, 1024–1042 (2020).

39. Chang, M. C., Hild, S. & Grieder, F. Nonhuman primate models for SARS-CoV-2 research: consider alternatives to macaques. *Lab Anim.* **50**, 113–114 (2021).

40. Kumari, P. et al. Neuroinvasion and encephalitis following intranasal inoculation of SARS-CoV-2 in K18-hACE2 mice. *Viruses* **13**, 132 (2021).

41. Peacock, T. P. et al. The altered entry pathway and antigenic distance of the SARS-CoV-2 Omicron variant map to separate domains of spike protein. Preprint at https://doi.org/10.1101/2021.12.31.474653 (2022).

42. Padmanabhan, P. & Dixit, N. M. Modelling how the altered usage of cell entry pathways by the SARS-CoV-2 Omicron variant may affect the efficacy and synergy of TMPRSS2 and cathepsin B/L inhibitors. Preprint at https://doi.org/10.1101/2022.01.13.476267 (2022).

43. Hui, K. P. Y. et al. SARS-CoV-2 Omicron variant replication in human bronchus and lung ex vivo. *Nature* **603**, 715–720 (2022).

44. Grau-Exposito, J. et al. Evaluation of SARS-CoV-2 entry, inflammation and new therapeutics in human lung tissue cells. *PLoS Pathog.* **18**, e1010171 (2022).

45. Zhao, H. et al. SARS-CoV-2 Omicron variant shows less efficient replication and fusion activity when compared with Delta variant in TMPRSS2-expressed cells. *Emerg. Microbes Infect.* **11**, 277–283 (2022).

46. Ricciardi, S. et al. The role of NSP6 in the biogenesis of the SARS-CoV-2 replication organelle. *Nature* **606**, 761–768 (2022).

47. Sun, X. et al. SARS-CoV-2 non-structural protein 6 triggers NLRP3-dependent pyroptosis by targeting ATP6AP1. *Cell Death Differ.* **29**, 1240–1254 (2022).

48. Cottam, E. M., Whelband, M. C. & Wileman, T. Coronavirus NSP6 restricts autophagosome expansion. *Autophagy* **10**, 1426–1441 (2014).

49. Benvenuto, D. et al. Evolutionary analysis of SARS-CoV-2: how mutation of non-structural protein 6 (NSP6) could affect viral autophagy. *J. Infect.* **81**, e24–e27 (2020).

50. Gosert, R., Kanjanahaluethai, A., Egger, D., Bienz, K. & Baker, S. C. RNA replication of mouse hepatitis virus takes place at double-membrane vesicles. *J. Virol* **76**, 3697–3708 (2002).

51. Liu, Z. et al. Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *Cell Host Microbe* **29**, 477–488 (2021).

52. Shah, M. & Woo, H. G. Omicron: a heavily mutated SARS-CoV-2 variant exhibits stronger binding to ACE2 and potently escapes approved CO VID-19 therapeutic antibodies. *Front. Immunol.* **12**, 830527 (2021).

53. Ye, G., Liu, B. & Li, F. Cryo-EM structure of a SARS-CoV-2 omicron spike protein ectodomain. *Nat. Commun.* **13**, 1214 (2022).

54. Carreno, J. M. et al. Activity of convalescent and vaccine serum against SARS-CoV-2 Omicron. *Nature* **602**, 682–688 (2022).

55. Torii, S. et al. Establishment of a reverse genetics system for SARS-CoV-2 using circular polymerase extension reaction. *Cell Rep.* **35**, 109014 (2021).

56. Chen, D. Y. et al. SARS-CoV-2 disrupts proximal elements in the JAK-STAT Pathway. *J. Virol.* **95**, e0086221 (2021).

57. Tzou, P. L., Tao, K., Pond, S. L. K. & Shafer, R. W. Coronavirus Resistance Database (CoV-RDB): SARS-CoV-2 susceptibility to monoclonal antibodies, convalescent plasma, and plasma from vaccinated persons. *PLoS ONE* **17**, e0261045 (2022).

58. Rambaut, A. et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **5**, 1403–1407 (2020).

59. Aksamentov I, R. C., Hodcroft, E. & Neher, R. Nextclade: clade assignment, mutation calling and quality control for viral genomes. *J. Open Source Softw.* **6**, 3773 (2021).

60. Jacob, A. et al. Derivation of self-renewing lung alveolar epithelial type II cells from human pluripotent stem cells. *Nat. Protoc.* **14**, 3303–3332 (2019).

Acknowledgements

We thank Y. Matsuura for providing plasmids containing fragments of the SARS-CoV-2 genome and the linker plasmid pMW-CoV-2-UTRlinker; the Department of Public Health, Massachusetts, for providing the clinical specimen containing Omicron virus; and the ICCB-Longwood Screening Facility of Harvard Medical School for assistance with immunofluorescence image acquisition and analysis. This work was supported by Boston University startup funds (to M.S. and F.D.); NIH S10 Shared Instrumentation Grants S10-OD026983 and SS10-OD030269 (to N.A.C.); a Peter Paul Career Development Award (to F.D.); and BMBF SenseCoV2 01KI20172A, DFG Fokus COVID-19, EN 423/7-1 and coronavirus research grants by the Bavarian State Ministry of Science and the Arts and Bavarian State Ministry of Health Bay-VOC (to A.E.). We thank the Evans Center for

Interdisciplinary Biomedical Research at Boston University School of Medicine for its support of the Affinity Research Collaborative on 'Respiratory Viruses: A Focus on COVID-19'.

Author information

Author notes

1. These authors contributed equally: Da-Yuan Chen, Chue Vin Chin

Authors and Affiliations

1. Department of Biochemistry, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, USA

   Da-Yuan Chen, Chue Vin Chin, Alexander H. Tavares, Nazimuddin Khan, Hasahn L. Conway & Mohsan Saeed

2. National Emerging Infectious Diseases Laboratories, Boston University, Boston, MA, USA

   Da-Yuan Chen, Chue Vin Chin, Devin Kenney, Alexander H. Tavares, Nazimuddin Khan, Hasahn L. Conway, Hans P. Gertje, Aoife K. O'Connell, Scott Adams, John H. Connor, Nicholas A. Crossland, Florian Douam & Mohsan Saeed

3. Department of Microbiology, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, USA

   Devin Kenney, Scott Adams, John H. Connor & Florian Douam

4. Florida Research and Innovation Center, Cleveland Clinic, Port St. Lucie, FL, USA

   GuanQun Liu & Michaela U. Gack

5. Brigham and Women's Hospital, Boston, MA, USA

Manish C. Choudhary & Jonathan Z. Li

6. Harvard Medical School, Cambridge, MA, USA

Manish C. Choudhary & Jonathan Z. Li

7. Center for Regenerative Medicine of Boston University and Boston Medical Center, Boston, MA, USA

Darrell N. Kotton

8. The Pulmonary Center and Department of Medicine, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, USA

Darrell N. Kotton & Markus Bosmann

9. Institute of Clinical and Molecular Virology, University Hospital Erlangen, Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany

Alexandra Herrmann & Armin Ensser

10. Department of Pathology and Laboratory Medicine, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, USA

Markus Bosmann, Yachana Kataria & Nicholas A. Crossland

11. Center for Thrombosis and Hemostasis, University Medical Center of the Johannes Gutenberg University, Mainz, Germany

Markus Bosmann

12. Department of Microbiology and Immunology, Stritch School of Medicine, Loyola University Chicago, Maywood, IL, USA

Susan C. Baker

13. Infectious Disease and Immunology Research Institute, Stritch School of Medicine, Loyola University Chicago, Maywood, IL, USA

Susan C. Baker

14. Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison, Madison, WI, USA

Robert N. Kirchdoerfer

Contributions

M.S. conceptualized the study. D.-Y.C., C.V.C., D.K., A.H.T., S.A., N.K., H.L.C., F.D. and M.S. performed experiments. G.L. and M.U.G. established and provided the modified CPER system. N.A.C., H.P.G. and A.K.O. performed histopathological and IHC analysis of mouse lungs. R.N.K. generated the 3D image of the S protein. M.C.C. and J.Z.L. performed sequencing analyses of recombinant viruses. D.N.K. provided iPS-cell-derived alveolospheres. S.C.B. and M.B. provided scientific input. A.H. and A.E. provided BACs containing the SARS-CoV-2 genome. J.H.C. provided the Omicron isolate. Y.K. provided plasma samples. MS wrote the manuscript, which was read, edited and approved by all authors.

Corresponding author

Correspondence to Mohsan Saeed.

Ethics declarations

Competing interests

R.N.K. is an inventor on US patent 10,960,070B2, entitled ‘Prefusion coronavirus spike proteins and their use’. M.S. is an inventor on a pending patent entitled ‘Immunogenic compositions and use thereof’ (USSN: 17/463,429). Y.K. received unrelated funding from Abbott Laboratories to assess the durability of SARS-CoV-2 antibodies for healthcare workers. M.B. received funds from ARCA Biopharma for unrelated COVID-19 research. F.D. received funding from Moderna, Flagship Pioneering and ARCA Biopharma for unrelated research. The remaining authors declare no competing interests.

Peer review

Peer review information

Nature thanks Ravindra Gupta and Olivier Schwartz for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Schematic overview of mutations in Omicron S.

Top, amino acid changes in Omicron S compared to the SARS-CoV-2 Wuhan-Hu-1 isolate (NCBI accession number: NC_045512). Numbering is based on Wuhan-Hu-1 sequence. Mutations not reported in previous variants of concern are shown in red. NTD, N-terminal domain; RBD, receptor-binding domain; RBM, receptor-binding motif. Bottom, location of Omicron changes on the trimeric S protein. Domains are coloured similarly in both panels.

Extended Data Fig. 2 Generating recombinant SARS-CoV-2 by CPER.

a, Schematic of recombinant SARS-CoV-2 generated by CPER (created with BioRender.com). S, spike; N, nucleocapsid. **b**, CPER protocol used in this study[24]. The SARS-CoV-2 genome was amplified into nine overlapping fragments. These fragments and a linker fragment (amplified from either pMW-CoV-2-UTRlinker or pGL-CPERlinker plasmid) were treated with PNK to phosphorylate 5' ends. The 5'-end- phosphorylated fragments were then stitched together by CPER, and the nicks in the resulting circular DNA

molecule were closed by treatment with DNA ligase. The CPER product was transfected into cells to rescue virus particles. **c**, ACE2/TMPRSS2/Caco-2 cells transfected with the SARS-CoV-2 CPER product were stained with an anti-nucleocapsid antibody on indicated days post-transfection. DAPI was used to stain the cell nuclei. NC, negative control generated by omitting fragment 9 from the CPER reaction. **d**, Virus titre in the culture medium of transfected cells at indicated days post-transfection, as measured by the plaque assay. The experiment was repeated twice. Individual values from both experiments are plotted.

Extended Data Fig. 3 Cytopathogenicity and replication of WT, Omi-S and Omicron in ACE2/TMPRSS2/Caco-2 cells.

a, Cell viability of SARS-CoV-2-infected ACE2/TMPRSS2/Caco-2 cells (MOI of 0.1) was quantified by the CellTiter-Glo assay at indicated times of infection. *P* values indicate a statistically significant difference between Omi-S and Omicron ($n = 3$ replicates). **b**, ACE2/TMPRSS2/Caco-2 cells were infected with WT, Omi-S, and Omicron at an MOI of 0.01, 0.1, or 1, and the viral titre in the culture medium was measured by the plaque assay at indicated times post-infection. Statistically significant difference between Omi-S and Omicron has been shown for the middle time point ($n = 3$ replicates). *p* values were calculated by a two-tailed, unpaired *t*-test with Welch's correction. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$; ns, not significant.

Extended Data Fig. 4 Comparison of Omicron isolate and recombinant Omicron in cell culture.

ACE2/TMPRSS2/Caco-2 (left) and Vero E6 (right) cells were infected with Omicron isolate or recombinant Omicron (generated by CPER) at an MOI of 0.01. The culture medium of infected cells was collected at indicated times, and the viral titre was measured by the plaque assay on ACE2/TMPRSS2/Caco-2 cells. The experiment was repeated twice, each time in three replicates. Error bars, mean ± s.d. ($n = 3$ replicates).

Extended Data Fig. 5 S-protein cleavage and fusogenicity of Omi-S and Omicron.

a, Western blot of S incorporated into virus particles. Virions generated in ACE2/TMPRSS2/Caco-2 cells were concentrated, and equal amount of total protein was loaded in each lane. S (antibody against S2 domain) and N were detected. Numbers at the bottom indicate mean ± s.d. of two independent experiments. **b**, S in infected cell lysates. ACE2/TMPRSS2/Caco-2 cells, infected at an MOI of 0.01, were collected at 24 hpi and processed for western blot with antibodies against S2 and nucleocapsid. β-actin served as an internal control. Numbers at the bottom indicate mean ± s.d. of two independent experiments. For gel source data, see Supplementary Fig. 2. **c**, Immunofluorescence staining of ACE2/TMPRSS2/293T cells with anti-nucleocapsid antibody. Nuclei was stained with DAPI. Infection was carried out at an MOI of 1 for 18 h. Left, representative images; right, size of 20 syncytia from two experimental repeats. *P* values were calculated by a two-tailed, unpaired *t*-test with Welch's correction.

Extended Data Fig. 6 Preferential cell entry of Omi-S and Omicron through the cathepsin-mediated pathway.

Vero E6 or ACE2/TMPRSS2/Vero E6 cells treated with E64d for 2 h were infected at an MOI of 0.5 and stained with anti-nucleocapsid antibody at 24 hpi for immunofluorescence analysis. Bar graphs show the percentage of infected cells relative to a DMSO control, for which the values were arbitrarily set at 100. Error bars, mean ± s.d. ($n = 4$ replicates). The experiment was performed only once. *P* values were calculated by a two-tailed, unpaired *t*-test with Welch's correction.

Extended Data Fig. 7 Clinical signs in Omi-S-infected mice.

a,**b**, K18-hACE2 mice ($n = 10$) inoculated intranasally with 1×10^4 PFU of Omi-S and described in Fig. 2a–c were monitored for body weight (**a**) and clinical score (**b**). Mice that lost 20% of their body weight (8 out of 10) were euthanized. The surviving mice did not show any signs of distress.

Source data

Extended Data Fig. 8 Lung pathology and nasal turbinate IHC in mice infected with WT, Omi-S and Omicron.

The lungs and nasal turbinates of K18-hACE2 mice intranasally inoculated with 1×10^4 PFU of WT, Omi-S and Omicron were collected at 2 dpi for histological analysis. **a**, Representative images of haematoxylin and eosin (H&E) staining for the detection of bronchiolar damage in the lungs of the infected mice. The bronchiolar epithelial necrosis is shown with arrows. **b**, IHC staining to detect SARS-CoV-2 S protein in the same area where bronchiolar necrosis was seen. No evidence of necrosis was found in bronchioles of mice infected with Omicron. Scale bars, 100 µm. **c**, Nucleocapsid distribution and abundance in the nasal passages of WT, Omi-S and Omicron. Left, representative images; Right, N immunoreactivity in nasal respiratory epithelium presented as percentage of the mean of WT (n = 2 mice for WT and Omicron; n = 3 mice for Omi-S). Chromogenic 3,3'-diaminobenzidine (DAB) IHC: (left panel, 12.5× total magnification; right panel, 400× total magnification). Higher-magnification views of nasal respiratory epithelium indicated by red squares are shown on the right. Scale bars: left, 800 µm; right, 20 µm.

Source data

Extended Data Fig. 9 Panel of chimeric viruses containing S and non-spike mutations.

a, Amino acid changes outside of S in Omicron BA.1 compared to D614-containing ancestral SARS-CoV-2. Proteins with amino acid changes are shown in red, whereas wild-type proteins are in blue. Amino acid numbers are according to individual viral proteins. **b**, Schematic representation of recombinant viruses generated by CPER. ORF7a was being replaced with mNeonGreen to generate reporter viruses. *Sequencing analysis showed that Omi-S/M lacked the A63T mutation in M, despite it being present in the plasmid used for CPER. S, spike; E, envelope; M, membrane; N, nucleocapsid.

Extended Data Fig. 10 Individual neutralization data.

Individual neutralization curves for the data presented in Fig. 4a,b are shown. The data represent the mean ± s.d. of three technical replicates. The curves were calculated based on a non-linear regression curve fit analysis in Prism. The dotted lines represent the limit of detection.

Extended Data Table 1 Information about serum samples

Supplementary information

Supplementary Figures

This file contains Supplementary Figs. 1 and 2, which show the gating strategy and the uncropped blots.

Reporting Summary

Source data

Source Data Fig. 2

Source Data Fig. 3

Source Data Extended Data Fig. 7

Source Data Extended Data Fig. 8

Rights and permissions

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Reprints and Permissions

About this article

Cite this article

Chen, DY., Chin, C.V., Kenney, D. *et al.* Spike and nsp6 are key determinants of SARS-CoV-2 Omicron BA.1 attenuation. *Nature* **615**, 143–150 (2023). https://doi.org/10.1038/s41586-023-05697-2

- Received: 16 March 2022
- Accepted: 04 January 2023
- Published: 11 January 2023
- Issue Date: 02 March 2023
- DOI: https://doi.org/10.1038/s41586-023-05697-2

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

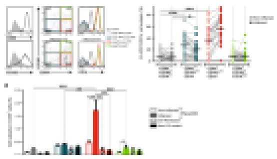
This article is cited by

- **Stricter US guidelines for 'gain-of-function' research are on the way — maybe**
    - Sara Reardon

Nature (2023)

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05697-2

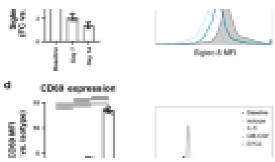
| Section menu | Main menu |



Two distinct colonic CD14+ subsets characterized by single-cell RNA profiling in Crohn's disease

22 January 2019

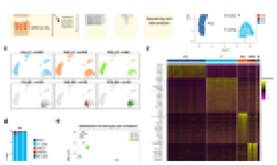
Laurence Chapuy, Marwa Bsat, … Marika Sarfati



Bidirectional crosstalk between eosinophils and esophageal epithelial cells regulates inflammatory and remodeling processes

10 May 2021

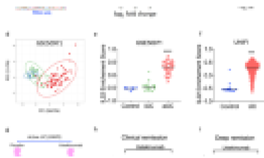
Julia L. M. Dunn, Julie M. Caldwell, … Marc E. Rothenberg



Single cell transcriptomic analysis of the immune cell compartment in the human small intestine and in Celiac disease

22 August 2022

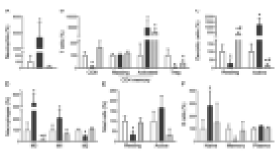
Nader Atlasy, Anna Bujko, … Hendrik G. Stunnenberg



Interleukin-22 regulates neutrophil recruitment in ulcerative colitis and is associated with resistance to ustekinumab therapy

03 October 2022

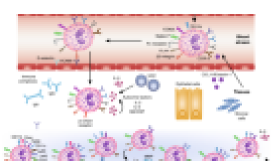
Polychronis Pavlidis, Anastasia Tsakmaki, … Nick Powell



Ulcerative colitis immune cell landscapes and differentially expressed gene signatures determine novel regulators and predict clinical response to biologic therapy

27 April 2021

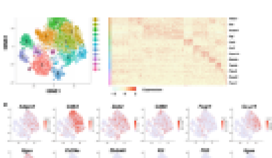
Harrison M. Penrose, Rida Iftikhar, … Suzana D. Savkovic



The emerging roles of eosinophils in mucosal homeostasis

10 March 2020

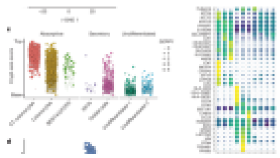
Kathleen Shah, Aline Ignacio, … Nicola. L. Harris



[Commensal microbiota drive the functional diversification of colon macrophages]

26 November 2019

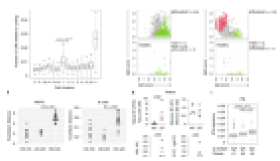
Byunghyun Kang, Luigi J. Alvarado, … Brian L. Kelsall



[Colonic epithelial cell diversity in health and inflammatory bowel disease]

27 February 2019

Kaushal Parikh, Agne Antanaviciute, … Alison Simmons



[Ulcerative colitis is characterized by a plasmablast-skewed humoral response associated with disease activity]

21 February 2022

Mathieu Uzzan, Jerome C. Martin, … Saurabh Mehandru

- Article
- Open Access
- [Published: 12 December 2022]

Active eosinophils regulate host defence and immune responses in colitis

- Alessandra Gurtner,
- Costanza Borrelli,
- Ignacio Gonzalez-Perez,
- Karsten Bach,
- Ilhan E. Acar,
- Nicolás G. Núñez,
- Daniel Crepaz,
- Kristina Handler,
- Vivian P. Vu,
- Atefeh Lafzi,
- Kristin Stirm,
- Deeksha Raju,
- Julia Gschwend,
- Konrad Basler,
- Christoph Schneider,
- Emma Slack,
- Tomas Valenta,
- Burkhard Becher,
- Philippe Krebs,
- Andreas E. Moor &
- …
- Isabelle C. Arnold

Show authors

Nature **volume 615**, pages 151–157 (2023)

- 10k Accesses

- 1 Citations

- 392 Altmetric
- Metrics details

Abstract

In the past decade, single-cell transcriptomics has helped to uncover new cell types and states and led to the construction of a cellular compendium of health and disease. Despite this progress, some difficult-to-sequence cells remain absent from tissue atlases. Eosinophils—elusive granulocytes that are implicated in a plethora of human pathologies[1,2,3,4,5]—are among these uncharted cell types. The heterogeneity of eosinophils and the gene programs that underpin their pleiotropic functions remain poorly understood. Here we provide a comprehensive single-cell transcriptomic profiling of mouse eosinophils. We identify an active and a basal population of intestinal eosinophils, which differ in their transcriptome, surface proteome and spatial localization. By means of a genome-wide CRISPR inhibition screen and functional assays, we reveal a mechanism by which interleukin-33 (IL-33) and interferon-γ (IFNγ) induce the accumulation of active eosinophils in the inflamed colon. Active eosinophils are endowed with bactericidal and T cell regulatory activity, and express the co-stimulatory molecules CD80 and PD-L1. Notably, active eosinophils are enriched in the lamina propria of a small cohort of patients with inflammatory bowel disease, and are closely associated with $CD4^+$ T cells. Our findings provide insights into the biology of eosinophils and highlight the crucial contribution of this cell type to intestinal homeostasis, immune regulation and host defence. Furthermore, we lay a framework for the characterization of eosinophils in human gastrointestinal diseases.

Main

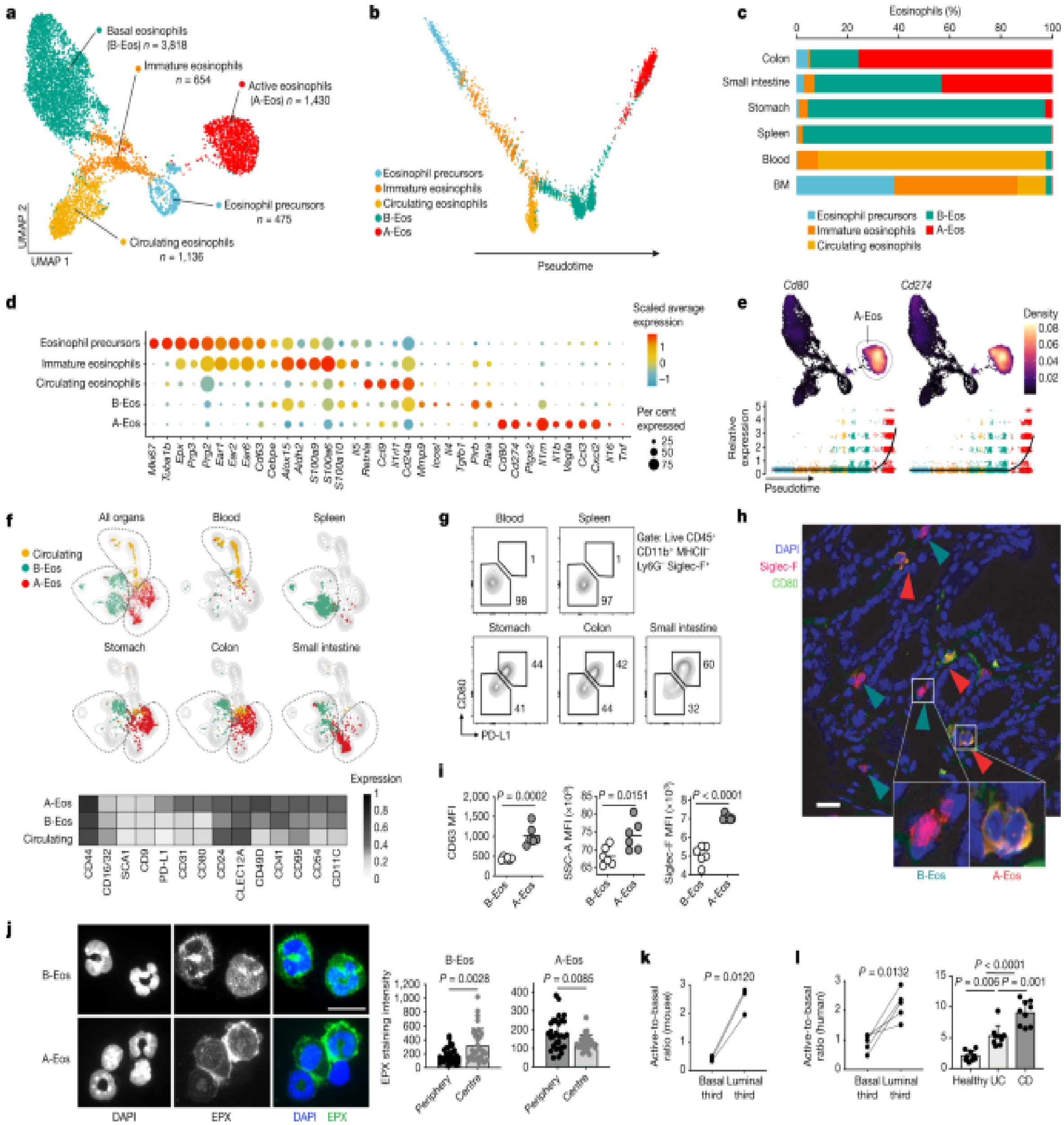
Eosinophils are granulocytes that reside mainly in the thymus, uterus, lung, adipose tissue and gastrointestinal (GI) tract[1]. Their accumulation is typical of disease states such as allergic airway inflammation, atopic dermatitis, eosinophilic oesophagitis and inflammatory bowel diseases (IBD)[2,3,4,5]. GI eosinophils contribute to various homeostatic processes, including

preserving the epithelial barrier, supporting tissue architecture, maintaining populations of immune cells and regulating local immune responses[6,7,8,9]. However, their function during intestinal inflammation is unclear[10]. Moreover, the presence of functionally distinct eosinophil subsets and their ontogenetic relationship have remained largely uninvestigated owing to technical challenges preventing their transcriptomic interrogation. Indeed, eosinophils are virtually absent from human and mouse single-cell RNA sequencing (scRNA-seq) atlases[11,12], and thus represent a blind spot in our understanding of cell-type-specific contributions to disease. Here, we fill this gap in knowledge by resolving eosinophil transcriptional and functional heterogeneity along their developmental trajectory from the bone marrow (BM) to tissues of residency, and by defining their role during intestinal inflammation.

A-Eos and B-Eos are two GI eosinophil subsets

By minimizing shear stress, degranulation and consequent transcript degradation (Extended Data Fig. 1a), we obtained single-cell transcriptomes from eosinophils isolated from the BM, blood, spleen, stomach, small intestine and colon of *Il5*-tg mice, a strain that has high eosinophil counts across tissues[13] (Extended Data Fig. 1b,c). We found that 89% of all cells widely expressed the bona fide eosinophil markers *Siglecf*, *Il5ra*, *Ccr3* and *Epx* (Extended Data Fig. 1d). Clustering revealed five subpopulations ordered along a developmental trajectory (Fig. 1a,b). Highly cycling precursors and immature eosinophils were primarily present in the BM, and circulating eosinophils were mainly in the blood. Two subsets, termed active eosinophils (A-Eos) and basal eosinophils (B-Eos), populated the GI tissues in varying proportions (Fig. 1c and Extended Data Fig. 1e).

Fig. 1: A-Eos and B-Eos are two distinct GI-resident eosinophil subsets.



a, Uniform manifold approximation and projection (UMAP) of eosinophil transcriptomes obtained from the BM, blood, spleen, small intestine, stomach and colon of *Il5*-tg mice ($n = 3$). **b**, Eosinophil differentiation trajectory. **c**, Subset distribution across organs (% of eosinophils). **d**, Expression of cluster marker genes. A complete list of cluster markers is available in Supplementary Table 1. **e**, Top, UMAP of *Cd80* and *Cd274* expression. Bottom, expression levels over pseudotime. **f**, Top, UMAP of eosinophil proteomic (spectral flow cytometry) profiles isolated from blood, spleen, stomach, colon and small intestine. Bottom, heat map of median

surface marker expression across subsets (*n* = 5, B6J). **g**, Representative FACS plots of A-Eos (PD-L1^{+}CD80^{+}) and PD-L1^{-}CD80^{-} eosinophils across organs. Numbers indicate percentage of eosinophils. **h**, Representative immunofluorescence of Siglec-F and CD80 in the mouse colon (*n* = 3, B6J). Arrows mark Siglec-F^{+}CD80^{+} A-Eos (red) and Siglec-F^{+}CD80^{-} B-Eos (green). Nuclei stained with DAPI. Scale bar, 20 µm. **i**, Mean fluorescence intensity (MFI) of CD63, SSC-A and Siglec-F in colonic A-Eos and B-Eos (*n* = 6, B6J). Medians are shown. Two-tailed unpaired Student's *t*-test. **j**, Left, representative images of cytospinned intestinal A-Eos and B-Eos stained with anti-EPX and DAPI (*n* = 3, *Il5*-tg). Scale bar, 10 µm. Right, quantification of EPX staining intensity at cell periphery and centre. Data are mean ± s.d. Two-tailed unpaired Student's *t*-test. **k**, Active-to-basal ratio in luminal versus basal third of colonic crypts (*n* = 3, B6J). Two-tailed paired Student's *t*-test. **l**, Left, active-to-basal ratio in luminal versus basal third of colonic crypts of healthy human colon cores (*n* = 5). Two-tailed paired Student's *t*-test. Right, active-to-basal ratio in samples from healthy individuals (5 individuals, 9 cores), patients with Crohn's disease (CD; 5 individuals, 9 cores) and patients with ulcerative colitis (UC; 4 individuals, 8 cores) samples. One-way ANOVA. Data are mean ± s.d. Patient information is provided in Supplementary Table 2. In **a**,**b**,**e**,**f**, dots represent single cells, coloured by cluster identity.

Eosinophil subsets exhibited distinct transcriptional profiles across organs and differed in their cytokine, effector-molecule and receptor repertoire, indicating that they have highly specialized functions (Fig. 1d and Extended Data Fig. 1f). Pseudotime analysis revealed that immature eosinophils downregulate stemness and proliferation programs, and transiently upregulate the expression of granular protein (*Epx*, *Prg2*, *Ear1*, *Ear2* and *Ear6*) and antimicrobial peptide genes (*S100a6*, *S100a9* and *S100a10*) (Extended Data Fig. 1g–i). Circulating eosinophils were characterized by high expression of *Retnla* and of the adhesion protein *Cd24a*, whereas B-Eos expressed effector molecules that are involved in tissue morphogenesis and remodelling, such as *Mmp9* and *Tgfb1* (Fig. 1d). Placed at the end of the differentiation trajectory, A-Eos were only found in organs of the GI tract and specifically expressed genes encoding multiple bioactive factors (*Il16*, *Tnf*, *Il1b*, *Ccl3*, *Cxcl2*, *Vegfa* and *Ptgs2*) and receptors (*Il1rn*, *Csf2rb*, *Tgfbr2*, *Ccr1*, *Cxcr4*, *Ptafr* and *Ahr*) (Fig. 1d and Extended Data Fig. 1j). Moreover,

their expression of the co-stimulatory molecules *Cd80* and *Cd274* (PD-L1) suggests that A-Eos are involved in immune modulation (Fig. 1e and Extended Data Fig. 1k). We thus focused our attention on this subset.

We profiled the surface proteome of blood, small intestine and colon eosinophils in B6J (wild-type) mice by spectral flow cytometry (fluorescence-activated cell sorting (FACS)) and found that the expression of PD-L1 and CD80 was sufficient to identify A-Eos (Fig. 1f–h and Extended Data Fig. 2a–c). PD-L1^{+}CD80^{+} cells expressed A-Eos markers at the protein and RNA levels (Extended Data Fig. 2d,e), and exhibited higher secretory activity[14,15,16] (CD63, CD9 and CD107a), granularity (SSC-A) and activation (Siglec-F) relative to B-Eos (Fig. 1i and Extended Data Fig. 2f). A-Eos also showed a peripheral distribution of eosinophil peroxidase (EPX), whereas granule localization in B-Eos, splenic and blood eosinophils was more cytosolic (Fig. 1j and Extended Data Fig. 2g). Of note, A-Eos and B-Eos differed in their spatial localization within the colonic mucosa, indicating exposure to and interactions with distinct cellular microenvironments: A-Eos were found significantly closer to the luminal extremity (luminal third), whereas B-Eos were retained near the submucosa (basal third) (Fig. 1k and Extended Data Fig. 2h,i). The presence of A-Eos was restricted to the GI tract, as PD-L1^{+}CD80^{+} eosinophils were not found by FACS (Extended Data Fig. 3a) or scRNA-seq (Extended Data Fig. 3b–e) in other tissues in which eosinophils reside, such as the uterus and adipose, and were only detected in small percentages in the thymus and peritoneum. A-Eos further differed from previously reported lung-resident populations and from inflammatory eosinophils recruited during a house dust mite (HDM) airway challenge[17] (Extended Data Fig. 3f,g).

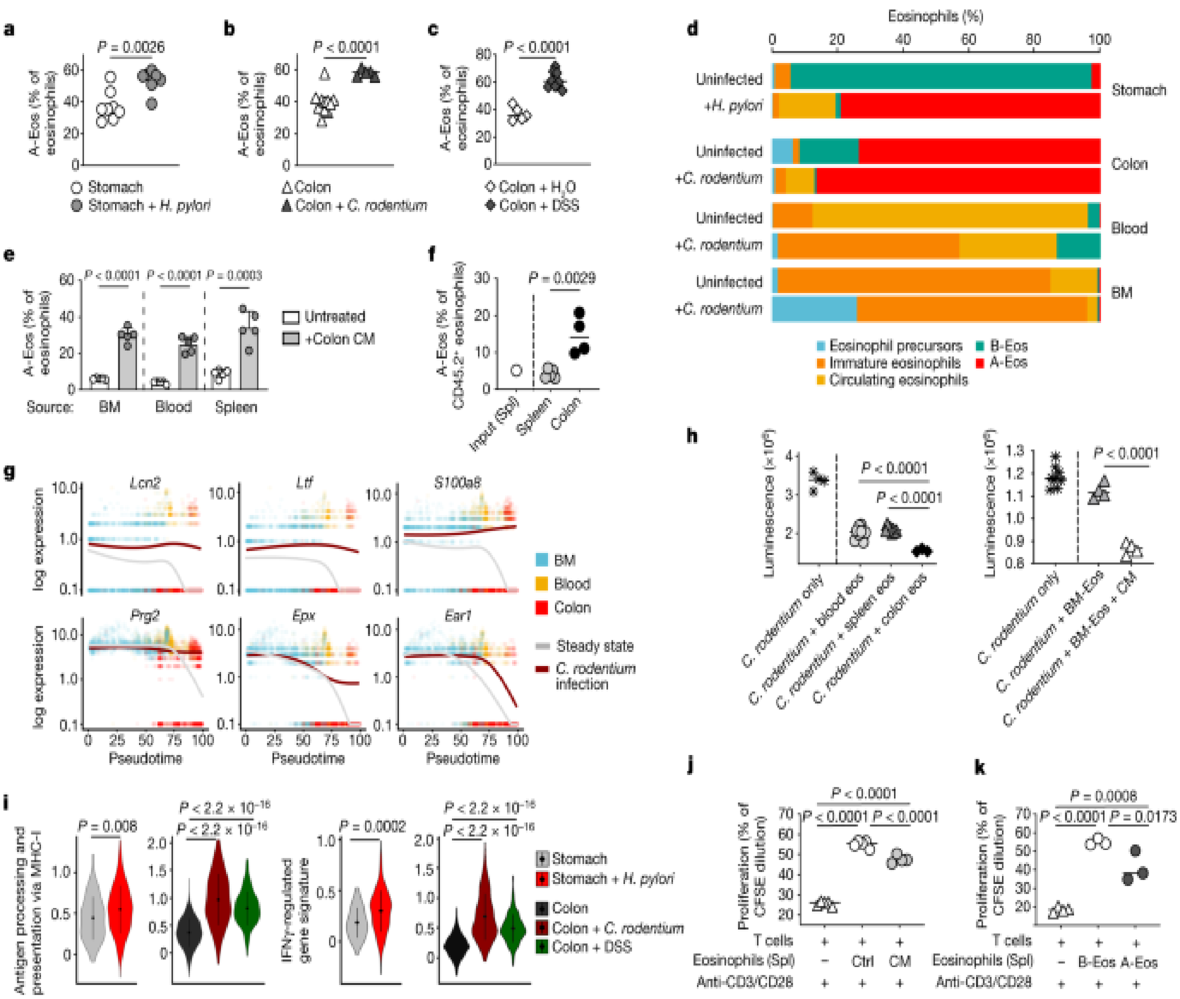
We next wondered whether A-Eos and B-Eos could also be found in the human GI tract, and whether their proportions are affected by colitis. We therefore subjected colon tissue microarrays (TMAs) from healthy individuals and from patients with IBD to major basic protein (MBP) and PD-L1 immunofluorescence analysis (Extended Data Fig. 3h). Similar to our observations in mice, MBP^{+}PD-L1^{+} A-Eos were found closer to the lumen than MBP^{+}PD-L1^{-} B-Eos, indicating phenotypic correspondence (Fig. 1l). Notably, the relative abundance of A-Eos (active-to-basal ratio) was twofold enriched in samples from patients with ulcerative colitis and fivefold

enriched in samples from patients with Crohn's disease, relative to healthy control individuals (Fig. 1l). This prompted us to investigate the role of A-Eos during intestinal inflammation.

A-Eos have antibacterial and regulatory functions

To assess how local insults affect the dynamics of the eosinophil subsets, we evaluated the frequency of PD-L1^{+}CD80^{+} A-Eos in three distinct experimental models of GI inflammation: acute *Citrobacter rodentium* infection in the colon, chronic *Helicobacter pylori* infection in the stomach and dextran sulfate sodium (DSS)-induced colitis (Fig. 2a–c and Extended Data Fig. 4a). A-Eos frequencies and numbers were significantly enriched across all models, reflecting what was observed in IBD and indicating that an increase in the active-to-basal ratio is a general response to epithelial damage and inflammation in the human and mouse gut.

Fig. 2: A-Eos have antibacterial and immune-regulatory functions.



a–**c**, A-Eos frequencies in *H. pylori*-infected (**a**; stomach, $n = 6$), *C. rodentium*-infected (**b**; colon, $n = 5$) and DSS-treated (**c**; colon, $n = 8$) mice relative to uninfected controls ($n = 5$–10, B6J). **a**,**b**, Data are pooled from two independent experiments. Medians are shown. Two-tailed unpaired Student's *t*-test. **d**, Percentage of eosinophil subsets across organs at steady state and during infection, as assessed by scRNA-seq. **e**, A-Eos frequencies after conditioning with colon CM. Input: BM-derived ($n = 5$, B6J), blood ($n = 5$, *Il5*-tg) and splenic ($n = 5$, *Il5*-tg) eosinophils. Data are mean ± s.d. Two-tailed unpaired Student's *t*-test. **f**, A-Eos frequencies among adoptively transferred CD45.2^{+} eosinophils in colon and spleen of host, 42 h after injection ($n = 4$, CD45.1). Input A-Eos frequency shown as a reference (spl, splenic eosinophils, $n = 2$, *Il5*-tg). Medians are shown. Two-tailed unpaired Student's *t*-test. **g**, Gene expression over common pseudotime at steady state (grey) and during *C. rodentium* infection (dark red). Dots indicate single cells, coloured by organ (BM, blood and colon). **h**, *C.*

rodentium (ICC180) viability after exposure to blood, splenic or colonic eosinophils ($n = 3$, pooled *Il5*-tg) or conditioned BM-Eos ($n = 3$, pooled B6J). Technical replicates and medians are shown. Two-tailed unpaired Student's *t*-test. **i**, Expression of MHC-I-restricted antigen processing and presentation signature and IFNγ-regulated genes. Genes used for scores and signatures are listed in Supplementary Table 3. Data are mean ± s.d. Two-sided Wilcoxon test ($n = 3$, *Il5*-tg). **j**,**k**, Proliferation of anti-CD3 and anti-CD28 (anti-CD3/CD28)-activated, carboxyfluorescein succinimidyl ester (CFSE)-labelled naive $CD4^+$ T cells co-cultured with conditioned splenic (**j**; Spl) or sorted GI (**k**) A-Eos and B-Eos ($n = 7$, *Il5*-tg mice). Medians are shown. One-way ANOVA.

To investigate the subset-specific transcriptional changes that occur during inflammation, we profiled eosinophils from the BM, blood and colon of *C. rodentium*-infected and from the stomach of *H. pylori*-infected *Il5*-tg mice by scRNA-seq (Extended Data Fig. 4b). We also retrieved eosinophil transcriptomes from an independent dataset of magnetically-enriched colonic $CD45^+$ cells of DSS-treated B6J (wild-type) mice[18] (Extended Data Fig. 4c). These single-cell profiles were integrated in the steady-state transcriptional embedding and mapped with high confidence to the existing clusters. Of note, merging the steady-state and the challenge datasets did not reveal novel inflammation-specific clusters (Extended Data Fig. 4d).

Infection strongly increased the active-to-basal ratio of eosinophils in the colon and stomach, and led to the accumulation of circulating eosinophils within infected tissues (Fig. 2d). Bacterial challenge further induced a relative expansion of immature eosinophils in the blood and BM. Core eosinophil populations are thus maintained during inflammation, but their proportions across organs vary to maximize the production of A-Eos at sites of infection. This compositional shift suggests alterations in the eosinophil differentiation path. Indeed, trajectory inference (Monocle; ref. [19]) and RNA velocity analysis (scvelo; ref. [20]) of BM, blood and colon eosinophils during *C. rodentium* infection placed A-Eos as originating directly from immature eosinophils—rather than from B-Eos, as observed at steady state (Extended Data Fig. 4e). Furthermore, circulating eosinophils found in the colon, but not in the blood, of *C. rodentium*-infected mice expressed multiple A-Eos markers, suggesting a bypassing of the B-Eos maturation stage and rapid

transition into A-Eos in situ (Extended Data Fig. 4f). Notably, single-cell fate probabilities computed with CellRank (ref. [21]) defined A-Eos as the major predicted terminal state for all eosinophil subsets, both at steady state and particularly during infection (Extended Data Fig. 4g). This suggests that B-Eos and circulating eosinophils are not alternative end states, but rather differentiation intermediates. In line with this, after in vitro exposure to colon supernatant (conditioned medium (CM)), eosinophils that differentiated from the BM (BM-Eos; mostly precursors and immature eosinophils), or that were derived from the blood (mainly circulating eosinophils) or spleen (mainly B-Eos), all equally acquired PD-L1^{+} and CD80^{+} surface expression in a dose-dependent manner, indicating that the potential of eosinophils to differentiate into A-Eos is maintained throughout their maturation (Fig. 2e and Extended Data Fig. 4h). We performed genetic fate mapping in $Id2^{CreERT2};Rosa^{EYPF}$ mice, a reporter strain in which Id2-Cre-expressing cells are inducibly labelled with EYFP. After a single tamoxifen pulse, the frequencies of colonic B-Eos among EYFP^{+} eosinophils decreased over time, whereas A-Eos frequencies increased, suggesting the conversion of B-Eos to A-Eos in vivo (Extended Data Fig. 4i). Similarly, adoptively transferred CD45.2 splenic eosinophils (B-Eos) migrated into the colon of CD45.1 hosts and showed evidence of in situ maturation into A-Eos (Fig. 2f). Cumulatively, these data suggest lineage plasticity and sequential ontogeny, with circulating eosinophils and B-Eos as metastable transition states along a dynamic differentiation continuum that culminates with A-Eos.

To investigate the transcriptional changes that are elicited by infection along the eosinophil maturation continuum, we aligned BM–blood–colon trajectories during steady state and *C. rodentium* infection to a common pseudotime axis[22]. At steady state, the expression of genes that encode granular proteins and antimicrobial peptides was only transiently upregulated by precursors and immature eosinophils, and therefore restricted to the BM; by contrast, infection induced the sustained expression of granulogenesis and antimicrobial gene programs in circulating and colonic A-Eos (Fig. 2g and Extended Data Fig. 4j,k). Notably, this did not result from altered recruitment kinetics, as assessed by 5-ethynyl-2'-deoxyuridine (EdU) pulsing, or from extramedullary haematopoiesis, as no lineage-committed progenitors (IL-5-Rα^{+}Lin^{-}Sca1^{-}CD34^{+}) were detected in the

colon afterinfection (Extended Data Fig. 4l,m). Moreover, the expression of CD63 in A-Eos was unaltered by bacterial challenge, indicating that the net increase in the levels of CD63 results from the accumulation of A-Eos rather than their enhanced secretory activity (Extended Data Fig. 4n). However, colonic A-Eos exhibited a marked change in morphology after *C. rodentium* infection, with evidence of cellular protrusions resembling extracellular DNA traps at sites of peripheral EPX accumulation (Extended Data Fig. 4o). We previously reported impaired bacterial clearance and enhanced colonic immunopathology in *C. rodentium*-infected eosinophil-deficient mice[23]. Hence, we assessed the bactericidal potential of A-Eos in co-culture with a bioluminescent *C. rodentium* strain. Colonic eosinophils (mainly A-Eos), as well as conditioned BM-Eos, exhibited significantly greater bactericidal activity with respect to blood (circulating), spleen (B-Eos) or unconditioned BM-Eos (immature eosinophils) (Fig. 2h). Our data therefore suggest that A-Eos are a highly specialized subset involved in bacterial control and endowed with antimicrobial and cytotoxic properties.

Across all our inflammation models, A-Eos specifically upregulated gene sets that are involved in immune modulation, IFNγ signalling and MHC-I-restricted antigen processing and presentation (Fig. 2i). Moreover, CellPhoneDB (ref. [24]) identified numerous potentially interacting ligand–receptor pairs between A-Eos, $CD4^+$ and $CD8^+$ T cells (Extended Data Fig. 5a). After treatment with DSS, eosinophil-deficient (PHIL) mice exhibited increased colitis severity (Extended Data Fig. 5b,c) and stronger T helper 17 cell (T_H17) responses relative to their wild-type littermates, as well as increased production of TNF and IFNγ by $CD4^+$ T cells[25] (Extended Data Fig. 5d). These data corroborate our previous report of an immune-regulatory role of eosinophils[23], which, given their relative abundance and specific expression of co-stimulatory molecules, may be attributed to A-Eos. Co-culture of both conditioned and unconditioned BM-Eos with OT-I $CD8^+$ T cells, but not OT-II $CD4^+$ T cells, resulted in robust T cell proliferation in an antigen-dependent manner, suggesting that eosinophils can present antigen via MHC-I and TCR interactions (Extended Data Fig. 5e). Conversely, conditioning of BM-Eos into A-Eos was required for the downregulation of $CD4^+$ T cell proliferation after anti-CD3- and anti-CD28-mediated stimulation (Fig. 2j). Indeed, only sorted intestinal A-Eos, and not B-Eos, were able to inhibit the proliferation of $CD4^+$ T cells (Fig. 2k),

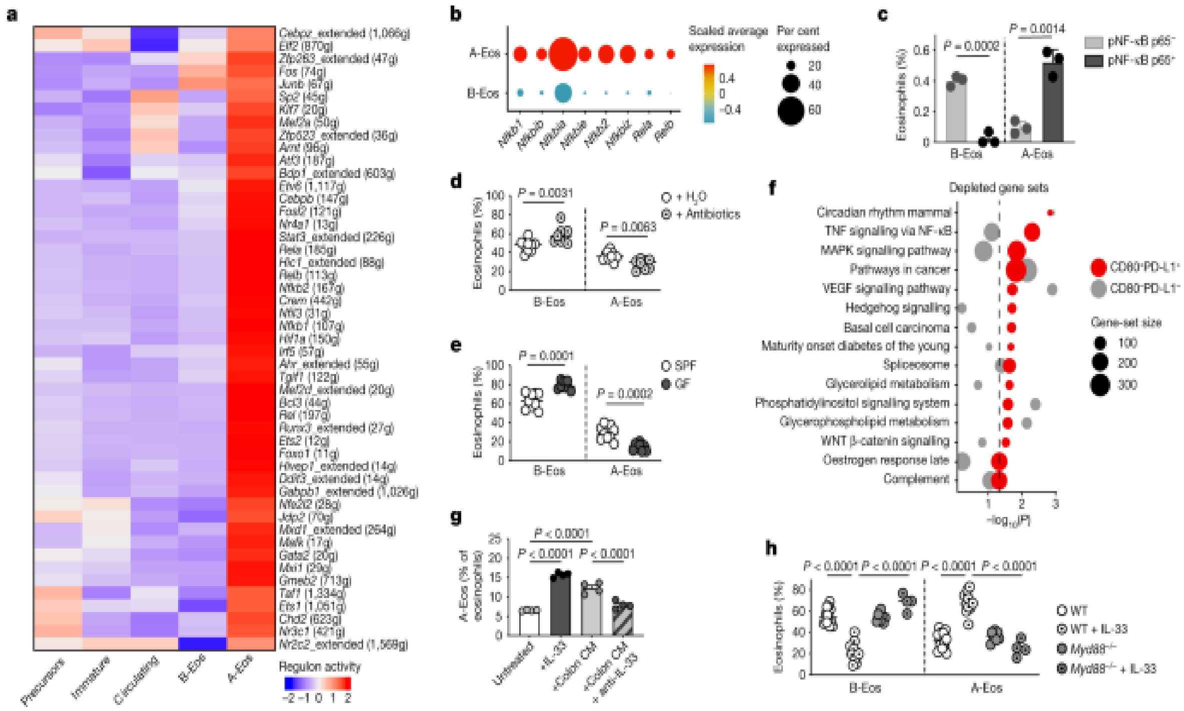
suggesting that this subset attenuates $CD4^+$ T cell responses during inflammation[9,23,26].

Of note, as IL-5 is a known driver of eosinophil maturation and survival[27], we conducted comparative flow cytometry and scRNA-seq analyses between B6J and *Il5*-tg mice. Aside from higher steady-state frequencies of A-Eos in *Il5*-tg mice (Extended Data Fig. 5f), we did not detect any transgene-specific effects during challenge (Extended Data Fig. 5g–i). Moreover, both subsets were similarly affected by anti-IL-5 treatment and equally depended on eotaxin–CCR3 interactions for their GI tissue accumulation (Extended Data Fig. 5j,k).

A-Eos maturation is induced locally by IL-33

Our data suggest that A-Eos have a dual antibacterial and immunomodulatory role during inflammation. We next sought to acquire a mechanistic understanding of the gene-regulatory network that governs the maturation, function and plasticity of A-Eos. Single-cell regulatory network inference and clustering (SCENIC; ref. [28]) revealed highly cluster-specific regulon activities and non-overlapping transcription factor profiles (Extended Data Fig. 6a,b). A-Eos exhibited high activity of several NF-κB-related regulons (*Rela*, *Relb*, *Nfkb1* and *Nfkb2*), which were predicted to directly govern the expression of *Cd274* and *Cd80* (Fig. 3a and Extended Data Fig. 6b). In line with the robust activation of this pathway that was indicated by both SCENIC and PROGENy analysis (Extended Data Fig. 6c), NF-κB signalling components were specifically upregulated in A-Eos and were expressed at significantly higher levels in colonic eosinophils compared with their blood and splenic counterparts (Fig 3b and Extended Data Fig. 6d). Furthermore, the co-localization of phosphorylated NF-κB p65 (pNF-κB p65) with $CD80^+$, but not $CD80^-$, eosinophils in the mouse colonic lamina propria indicates selective activation of canonical NF-κB signalling in A-Eos (Fig 3c and Extended Data Fig. 6e). Notably, NF-κB inhibition in vitro abolished BM-Eos conditioning into A-Eos (Extended Data Fig. 6f).

Fig. 3: A-Eos maturation is induced locally by IL-33.



a, Activity of A-Eos-specific regulons across clusters. g, number of genes in regulon. **b**, Expression of NF-κB signalling components. **c**, Quantification of pNF-κB $p65^{+}$ cells in colonic A-Eos and B-Eos ($n = 3$, B6J). Data are mean ± s.e.m. Two-tailed unpaired Student's *t*-test. **d**,**e**, A-Eos and B-Eos frequencies in antibiotic-treated (**d**) ($n = 16$, B6J) and germ-free (GF) (**e**) ($n = 9$, B6J) mice relative to controls. SPF, specific pathogen free. **d**, Data are pooled from two independent experiments. Medians are shown. Two-tailed unpaired Student's *t*-test. **f**, Depleted gene sets in PD-L1^{+}CD80^{+} A-Eos (red) and PD-L1^{-}CD80^{-} eosinophils (grey), relative to BM stem cells. Kolmogorov–Smirnov test. Dot size indicates gene-set size. Dashed line indicates $P = 0.05$. **g**, A-Eos frequencies after conditioning of BM-Eos with IL-33, colon CM and anti-IL-33 ($n = 2$, pooled B6J). Technical replicates and mean ± s.e.m. are shown. One-way ANOVA. **h**, Colonic A-Eos and B-Eos frequencies in B6J ($n = 21$) and *Myd88*$^{-/-}$ ($n = 15$) mice treated with IL-33, relative to untreated controls. Medians are shown. Two-tailed unpaired Student's *t*-test. WT, wild type.

Owing to their proximity to the lumen, we speculated that A-Eos might be induced by microbiota-derived cues signalling through the TLR–NF-κB pathway. Indeed, the proportion of A-Eos in the colon was significantly reduced after the depletion of commensal bacteria by broad-spectrum

antibiotics (Fig. 3d and Extended Data Fig. 6g) as well as in germ-free mice (Fig. 3e and Extended Data Fig. 6h). Germ-free mice also exhibited a marked reduction in eosinophil secretory activity, most prominently in A-Eos (Extended Data Fig. 6h). However, A-Eos frequencies were not affected by TLR2 or TLR4 deficiency (Extended Data Fig. 6i), suggesting independence from these major bacterial recognition pathways.

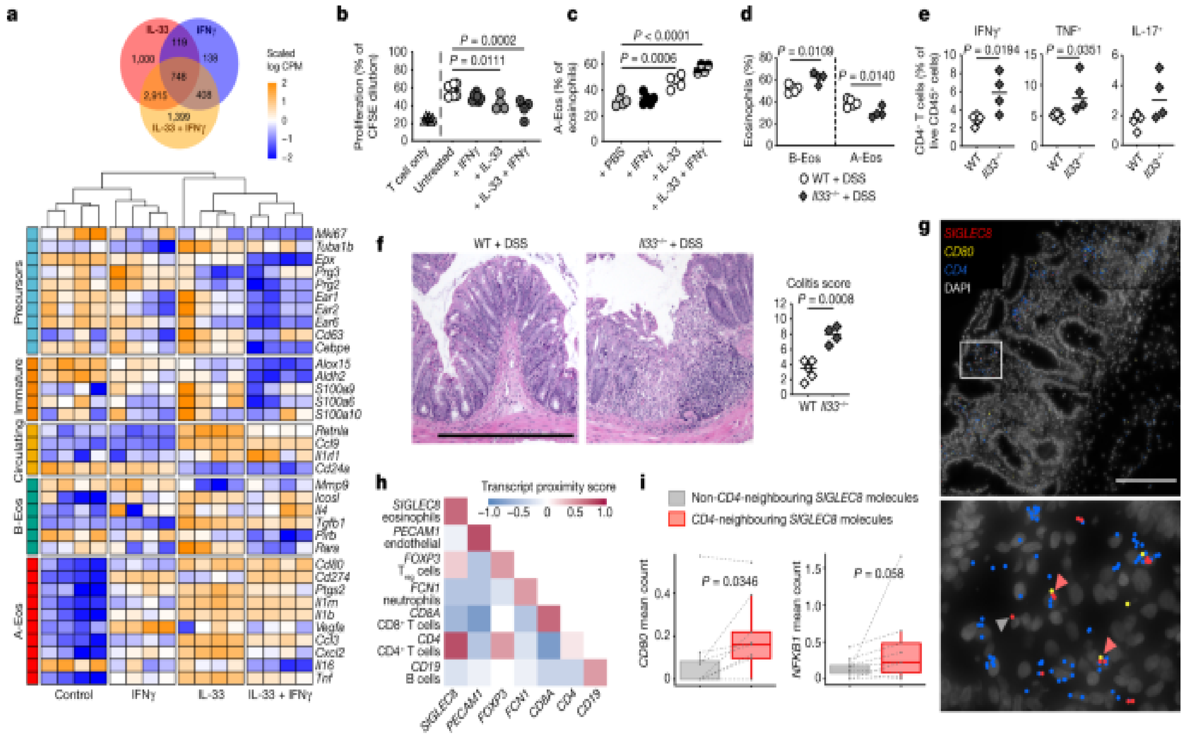
To identify regulatory checkpoints of A-Eos differentiation, we conducted an in vitro genome-wide CRISPR inhibition screen (Extended Data Fig. 7a). We found that single guide RNAs (sgRNAs) targeting genes involved in NF-κB and MAPK signalling were significantly depleted in PD-L1^{+}CD80^{+} but not in PD-L1$^{−}$CD80$^{−}$ eosinophils, compared to BM stem cells (Fig. 3f and Extended Data Fig. 7b). This observation is in line with our transcriptome analysis (Extended Data Figs. 1f and 6c) and suggests that activation of these pathways is required for A-Eos maturation. Notably, in vitro stimulation with the alarmin IL-33—but not with other cytokines such as IL-22, IL-25 and TNF, the levels of which increase during inflammation (Extended Data Fig. 7c,d)—was sufficient to induce A-Eos marker expression in a dose-dependent manner (Extended Data Fig. 7e,f). Moreover, IL-33 neutralization significantly reduced the differentiation of conditioned BM-Eos into A-Eos (Fig. 3g). Treating BM-Eos with IL-33 quickly led to the phosphorylation of p38 and p65, induced the expression of *Cd274*, *Cd80* and several other A-Eos markers, and further upregulated the surface presentation of the IL-33 receptor ST2 (Extended Data Fig. 7g–i). In vivo, ST2 was expressed at higher levels by A-Eos than B-Eos, suggesting a positive feedback loop to promote tissue adaptation (Extended Data Fig. 7j). Of note, we did not detect ST2 expression in lung, adipose, uterine, peritoneal or thymic eosinophils, which further suggests that the induction of A-Eos by IL-33 may be specific to the GI tract in homeostatic conditions (Extended Data Fig. 7k). IL-33 is known to activate the p38–MAPK and NF-κB pathways via the ST2–MyD88 signalling axis[29]. Indeed, ST2 deficiency abolished the effects of IL-33 treatment in BM-Eos, and significantly reduced their ability to be conditioned by colon CM (Extended Data Fig. 7l). In vivo, treatment with recombinant IL-33 markedly increased the frequencies of A-Eos in the colon and other organs in a MyD88-dependent manner (Fig. 3h and Extended Data Fig. 7m). Finally, A-Eos frequencies at steady state were reduced in the small intestine and stomach

of *Il33*$^{-/-}$ mice, but not in the colon, indicating that alternative, possibly microbiota-dependent mechanisms may contribute to A-Eos differentiation in the healthy colon (Extended Data Fig. 7n).

Role of IFNγ in A-Eos regulatory functions in colitis

The analysis of our challenge dataset by SCENIC suggests that signalling downstream of IFNγ is increased during inflammation in A-Eos. In particular, *C. rodentium* infection shifted the regulatory landscape towards signalling through STATS (*Stat1*, *Stat3*, *Stat4*, *Stat5b* and *Stat6*) and IRFs (*Irf1*, *Irf2*, *Irf5*, *Irf7* and *Irf9*) (Extended Data Fig. 8a). Notably, *Ifngr1* expression was restricted to the A-Eos subset (Extended Data Fig. 1j), and its deficiency in the eosinophil compartment results in decreased *C. rodentium* clearance and deregulated T cell responses during *H. pylori* infection[23]. To analyse the interplay of IL-33 and IFNγ in regulating A-Eos functions, we performed bulk RNA-seq of BM-Eos treated with IL-33, IFNγ or a combination thereof (Extended Data Fig. 8b). Treatment with IL-33 induced NF-κB signalling and the expression of A-Eos markers, whereas IFNγ treatment strongly upregulated the expression of *Cd274* and genes involved in antigen presentation (Fig. 4a and Extended Data Fig. 8c). Functionally, IL-33 and IFNγ treatment endowed BM-Eos with an increased ability to downregulate the proliferation of $CD4^+$ T cells (Fig. 4b). Notably, the synergistic effect of IL-33 and IFNγ not only increased the levels of A-Eos in vitro (Extended Data Fig. 8d), but also shifted the transcriptome of BM-Eos to a more mature state by downregulating granular protein and antimicrobial genes (Fig. 4a and Extended Data Fig. 8c). Treating A-Eos with IFNγ further induced granule mobilization and focal aggregation (Extended Data Fig. 8e). These results suggest a negative feedback loop on the synthesis of granular proteins and antimicrobial peptides, with their release being induced and their transcription being repressed by IFNγ signalling[23,30].

Fig. 4: A-Eos co-localize with $CD4^+$ T cells in patients with IBD.



a, Top, Venn diagram of significant differentially expressed genes (DEGs) (false discovery rate (FDR) < 0.05, logFC > |2|) in BM-Eos treated with IL-33 and/or IFNγ (*n* = 4, B6J). All DEGs are listed in Supplementary Table 4. Bottom, expression of subset markers across conditions. Columns are clustered, rows are scaled. CPM, counts per million. **b**, Proliferation of anti-CD3/CD28-activated, CFSE-labelled naive $CD4^+$ T cells co-cultured with BM-Eos conditioned with IL-33 and/or IFNγ (*n* = 4, B6J). Data are pooled from two independent experiments. Medians are shown. One-way ANOVA. **c**, A-Eos frequencies in mice treated with IL-33 and/or IFNγ (*n* = 5, B6J). Medians are shown. One-way ANOVA. **d**, A-Eos and B-Eos frequencies in DSS-treated B6J (*n* = 5) and $Il33^{-/-}$ (n = 4) mice. Medians are shown. Two-tailed unpaired Student's *t*-test. **e**, Frequencies of IFNγ-, IL-17- and TNF-expressing colonic $CD4^+$ T cells from mice in **d**. Medians are shown. Two-tailed unpaired Student's *t*-test. **f**, Left, representative haematoxylin and eosin (H&E)-stained colonic sections of mice in **d**. Scale bar, 100 µm. Right, colitis score assessed by histopathological examination. Medians are shown. Two-tailed unpaired Student's *t*-test. **g**, Representative molecular cartography images of human ulcerative colitis samples. Nuclei are stained with DAPI; *CD4*, *SIGLEC8* and *CD80* RNA molecules are shown in blue, red and yellow, respectively. Scale bar, 200 µm. **h**, Pairwise proximity score

of transcripts across slides. The score indicates the fraction of slides in which the proximity of a pair of transcripts is significantly higher than expected by chance. *P* values are computed using a permutation test (Methods). T_{reg} cells, regulatory T cells. **i**, Mean counts per slide of *CD80* and *NFKB1* transcripts in the proximity (<10 µm) of *SIGLEC8* transcripts spatially associated with *CD4* molecules versus *SIGLEC8* molecules not associated with *CD4* molecules. The central line in the box plot represents the median count per slide, the lower and upper hinge correspond to the first quartiles and the whisker extends from the hinge to the smallest or largest value no further than 1.5 times the interquartile range (IQR) from the hinge. Two-sided paired Wilcoxon test (17 regions of interest (ROIs), $n = 4$ patients).

In vivo, treatment with IFNγ potentiated the effects of IL-33, increasing colonic A-Eos frequencies to the levels observed during colitis (Fig. 4c). Consistently, ablation of IFNGR signalling in the eosinophil compartment impaired the upregulation of PD-L1 in response to infection (Extended Data Fig. 8f). Notably, IFNγR neutralization abrogated *Cd274* expression but did not affect the presence of the A-Eos subset in the steady-state colon, as assessed by scRNA-seq (Extended Data Fig. 8g). In line with our in vitro results, this treatment led to the upregulation of granular protein and antimicrobial peptide genes in A-Eos (Extended Data Fig. 8g).

Our data indicate that IFNγ potentiates but it is not sufficient to induce the A-Eos phenotype, which instead relies on IL-33 signalling. Indeed, IL-33-deficiency prevented the colonic accumulation of A-Eos after treatment with DSS (Fig. 4d). *Il33*$^{-/-}$ mice also suffered from increased DSS-induced colitis and did not restrict effector T cell responses (Fig. 4e,f), thus phenocopying eosinophil deficiency. Cumulatively, our data suggest that IL-33 promotes the accumulation of A-Eos during colitis, and that A-Eos limit pathogen incursions and prevent excessive tissue damage through their bactericidal and T-cell-regulatory activities.

A-Eos co-localize with CD4^{+} T cells in patients with IBD

Finally, we performed multiplexed in situ RNA imaging (molecular cartography) in colon sections from individuals with ulcerative colitis (A.L. et al., manuscript in preparation) and found that *CD4* transcripts significantly co-localized with *SIGLEC8* transcripts (Fig. 4g,h and Extended Data Fig. 8h). Indeed, 55% of *SIGLEC8*$^+$ segmented areas were also positive for *CD4* (Extended Data Fig. 8i), suggesting that eosinophils and CD4$^+$ T cells are in close spatial proximity in the colon of patients with IBD. In a segmentation-free approach, *CD4*-neighbouring *SIGLEC8* RNA molecules were significantly more associated with the A-Eos markers *CD80*, *VEGFA* and *CSF2RB* than were non-*CD4*-neighbouring *SIGLEC8* molecules, indicating that CD4$^+$ T cells preferentially interact with A-Eos (Fig. 4i and Extended Data Fig. 8j). *CD4*-neighbouring *SIGLEC8* RNA molecules were also significantly associated with NF-κB (*NFKB1*) and IFNγ (*IFNGR1*, *STAT1* and *IRF1*) signalling components (Fig. 4i and Extended Data Fig. 8j), which indicates that the same pathways might drive interactions between A-Eos and CD4$^+$ T cells in mouse and human colitis.

Discussion

Neutralizing antibodies against the cytokine IL-5 are widely used in severe eosinophilic asthma to inhibit the differentiation of eosinophils[31]. Whether intestinal eosinophils can be exploited as therapeutic targets in IBD is still unknown, and thus a thorough investigation of their functions in the healthy and inflamed gut is warranted. Owing to the technical challenges involved in profiling these elusive cells, eosinophils have long remained overlooked in colitis. Here, we identify a subset of GI-resident eosinophils that are enriched in patients with IBD and in experimental models of colitis. In response to bacterial cues, IL-33 and IFNγ signalling, A-Eos exert a protective role on the intestinal mucosa by means of their antibacterial and immunomodulatory activity. Impaired accumulation of A-Eos in the inflamed colon worsens inflammation and leads to the hyperactivation of CD4$^+$ T cells during acute colitis. However, the strong effector and cytotoxic potential of this subset can potentially also favour tissue damage in settings of chronic inflammation. More studies are needed to elucidate the extent and nature of their protective activities in human intestinal homeostasis and inflammation, and whether these can be targeted for the treatment of IBD.

Methods

Mice

All experiments were performed on 6–16-week-old mice. Mice of a given strain were randomly divided into the different groups and both males and females were included in studies. Treatments and study termination was performed by two or more experimenters and blinding during result assessment was done by converting animal identifiers into numbers during sample processing and analysis. C57BL/6J (B6J, stock no. 000664) and dCas9-KRAB (stock no.030000) mice were obtained from The Jackson Laboratory; OT-1 (stock no. 003831), OT-II (stock no. 004194), *MyD88*$^{-/-}$ (ref. [32]), *Tlr2*$^{-/-}$ (stock no. 004650), CD45.1 (stock no. 002014) and *Tlr4*$^{-/-}$ (ref. [33]) mice were obtained from a local live mouse repository. *Id2*^{CreERT2}*;Rosa26*^{EYFP} mice[34], *Il5*-transgenic mice[35] and *Ifngr2*$^{fl/fl}$ mice[36] have been previously described. *Il33*$^{-/-}$ (ref. [37]) were obtained through the RIKEN Center for Developmental Biology (accession number CDB0631K) and *St2*$^{-/-}$(ref. [38]) mice have been described and were backcrossed onto a C57BL/6J background. Eosinophil-deficient mice[39] (PHIL) and mice expressing Cre under the EPX promoter[40] (Eo-*Cre*) were obtained from J. J. Lee. Chow and water were available ad libitum, unless specified. All mice were in the B6J background and maintained on a 12-h light–12-h darkness schedule. Mice were housed and bred under specific-pathogen-free conditions in accredited animal facilities. Germ-free mice were bred and maintained in open-top cages within flexible-film isolators, supplied with HEPA-filtered air, and autoclaved food and water ad libitum. At the experimental end-point, mice were euthanized by increasing CO_2 concentrations. All experimental procedures at the University of Zurich and Bern were performed in accordance with Swiss Federal regulations and approved by the Cantonal Veterinary Office and/or in accordance with the European Communities Council Directive (86/609/EEC), Czech national guidelines, institutional guidelines of the Institute of Molecular Genetics and approved by the Animal Care Committee.

Animal experiments

Antibody neutralization

For the 10-day treatment: 7–8-week-old female and male mice (B6J) were injected intraperitoneally twice a week with 0.5 mg anti-IL-5 (BE0198 BioXCell, TREK5) or anti-keyhole limpet haemocyanin isotype control (BE0090, BioXCell, LTF-2), or anti-IFNγR (BE0029, BioXCell, GR-20) or anti-CCR3 (BE0316 clone 6S2-19-49) or anti-horseradish peroxidase isotype control (BE0088, BioXCell, HRPN) antibodies for 10 days before the study end-point.

Depletion of intestinal commensals by antibiotic treatment

Seven-to-eight-week-old female mice (B6J) were treated for 10 consecutive days with ampicillin (1 g l^{-1}; A0166 Sigma), vancomycin (500 mg l^{-1}; A1839,0001 Applichem), neomycin sulfate (1 g l^{-1}; 4801 Applichem), and metronidazole (1 g l^{-1}; H60258 Alfa Aesar) in autoclaved drinking water, as previously described[41]. Water bottles were monitored and refilled twice per week.

Adoptive transfer

A total of 10^6 magnetically selected splenic eosinophils of 6–12-week-old *Il5*-tg female and male mice were injected intravenously in 100 µl PBS into CD45.1 recipients (8–12-week-old female and male mice). Organs were collected 42 h after injection.

DSS-induced colitis

Six-to-twelve-week-old female and male mice (PHIL, B6J and *Il33*$^{-/-}$) were treated with 2.5% DSS (w/v; 9011-18-1, MP Biomedicals) dissolved in autoclaved drinking water for five days, followed by three days of regular water, before organs were collected. Water bottles were monitored and refilled twice per week.

H. pylori challenge

Six-to-twelve-week-old female and male mice (*Il5*-tg and B6J) were infected orally with the *H. pylori* strain PMSS1 (10^7 colony-forming units, CFU) and analysed four weeks after infection. The PMSS1 strain, a clinical isolate of a patient with a duodenal ulcer, was grown on horse blood agar plates followed by liquid culture, as previously described[23]. Cultures were routinely assessed by light microscopy for contamination, morphology and motility. *C. rodentium:* 6–12-week-old female and male mice (*Il5*-tg and B6J) were infected orally with the nalidixic-acid-resistant *C. rodentium* strain ICC169 (ATCC 51549, 10^8 CFU) and analysed 13 days after infection. Bioluminescent *C. rodentium* strains ICC180 (ICC169 derivative, nalidixic-acid- and kanamycin-resistant) was a gift from G. M. Frankel and was previously described[42]. Both strains were grown on agar plates (1.5%; A0927 Applichem), followed by single-colony picking and overnight culture in antibiotic-supplemented Luria broth (nalidixic acid, 50 μg ml^{-1}; N4382 Sigma and/or kanamycin, 50 μg ml^{-1}; 420311 Sigma).

Cytokine administration

Seven-to-eight-week-old female mice (*MyD88*$^{-/-}$ and B6J) were injected intraperitoneally every other day with three total doses of 0.5 mg rec-IL-33 (210-33, PeproTech) and/or IFNγ (315-05, PeproTech) or with PBS control.

EdU labelling

Seven-to-eight-week-old female and male mice (*Il5*-tg and B6J) were infected orally with *C. rodentium* or left uninfected; four days before analysis mice were injected with EdU (2.5 mg per mouse, 900584 Sigma).

HDM challenge

Four-month-old female mice (B6J) received 1 μg HDM extract in 50 μl PBS intratracheally for sensitization (day 0) and were then challenged once a day with 10 μg HDM in 50 μl PBS for 5 days (day 7–11). Lungs were collected 14 days after the sensitization.

Tamoxifen injection

Six-to-twelve-week-old female and male mice (*Id2^{CreERT2};Rosa26^{EYFP}*) were gavaged with a single dose of tamoxifen (T5648 Sigma). Tamoxifen was dissolved in a small volume of 100% ethanol (pre-warmed at 50 °C) and then resuspended in corn oil (pre-warmed at 50 °C) to the final concentration of 5 mg per mouse. Organs were collected 2 h, 2 and 4 days after the injection.

Preparation of single-cell suspensions from tissues

GI tissues

Stomach, colon and small intestine were collected, cleaned of faecal matter and cut longitudinally. Organs were washed in PSB and cut into pieces (1–2 cm) and Peyer's patches were removed from the small intestine. Pieces were washed twice in a shaking incubator with wash buffer (2% BSA, 100 U ml^{-1} penicillin–streptomycin and 5 mM EDTA in HBSS, 25 min, 37 °C). Tissues were then rinsed in cold PBS and digested for 50 min at 37 °C in complete medium (10% FBS and 100 U ml^{-1} penicillin–streptomycin (P0781 Sigma) in RPMI-1640) containing 15 mM HEPES (H0887 Sigma), 0.05 mg ml^{-1} DNase I (10104159001 Roche) and an equal amount of 250 U ml^{-1} type IV (C5138 Sigma) and type VIII collagenase (C2139 Sigma) (for colon and small intestine), or 500 U ml^{-1} type IV collagenase (C5138 Sigma) (for stomach). Cells were passed through a 70-µm cell strainer, centrifuged for 8 min and layered onto a 40/80% Percoll (17089101 Cytiva) gradient (18 min, 2,100*g*, 20 °C, no brake). The interphase was collected and washed in PBS.

Lung

Lungs were perfused with PBS, collected and cut into pieces before digestion in complete medium supplemented with 500 U ml^{-1} type IV collagenase (Sigma) and 0.05 mg ml^{-1} DNase I (Roche) for 50 min at 37 °C. Lungs were then passed through a 70-µm cell strainer and mesh with syringe plungers. To reduce macrophage contamination (Siglec-F^{+}), cells were plated in complete RPMI medium for 1 h at 37 °C.

Blood

Blood was sampled by post mortem cardiac puncture in 2% BSA 5 mM EDTA PBS. For *Il5*-tg mice, the suspension was layered over Histopaque 1119 (density of 1.119 g ml^{-1}; 11191 Sigma-Aldrich) and centrifuged at 800*g* for 20 min and the interphase was washed in PBS. Red blood cells were lysed in ice-cold distilled water for 30 s.

Bone marrow

Femur and tibia were flushed using complete RPMI medium and a 23-gauge needle. The content was collected, filtered through a 40-µm cell strainer and red blood cells were lysed in ice-cold distilled water for 30 s.

Spleen, lymph nodes and thymus

Spleen and lymph nodes were collected and meshed through a 40-µm cell strainer using a syringe plunger, and red blood cells were lysed in ice-cold distilled water for 30 s.

Peritoneal fluid

Peritoneal cavity was perfused with 5 ml PBS with a 21-gauge needle and the inflated area was massaged for 30 s, to disperse the solution. The peritoneal liquid was collected and cells were plated in complete RPMI medium for 1 h at 37 °C to remove adherent cells.

Adipose tissue

Lungs were perfused with PBS and the perigonadal adipose depot was isolated, removing any visible gonadal tissue. The tissue was minced into small pieces and digested in complete DMEM medium supplemented with 0.2 mg ml^{-1} Liberase (05401020001 Roche) and 0.05 mg ml^{-1} DNase I (Roche) for 50 min at 37 °C. Suspensions were filtered through a 100-µm cell strainer and centrifuged at 1,000*g* for 10 min. The pellet was collected and washed in PBS.

Uterus

Uterus was collected, cut longitudinally and washed in PBS. Pieces were shaken in wash buffer (2% BSA, 100 U ml^{-1} penicillin–streptomycin and 5 mM EDTA in HBSS, 25 min, 37 °C). The tissue was then rinsed in cold PBS and digested for 50 min at 37 °C in complete medium containing 0.05 mg ml^{-1} DNase I (Roche) and 0.2 mg ml^{-1} Liberase (Roche). Cells were passed through a 70-μm cell strainer, centrifuged and washed in PBS.

Unless specified, all centrifugation steps were performed at 500*g* for 8 min at 10 °C.

Magnetic cell enrichment

Eosinophils of 6–12-week-old female and male mice (*Il5*-tg) were positively enriched using a PE anti-mouse Siglec-F antibody (562068 BD Biosciences; E50-2440) and anti-PE microbeads (130-042-401 Miltenyi Biotech), according to the manufacturer's instructions. Immune cells of 7–8-week-old female mice (B6J) were positively enriched using anti-CD45 microbeads (130-052-301 Miltenyi Biotech), according to the manufacturer's instructions.

scRNA-seq

Single-cell capture and library preparation

Whole-transcriptome analyses of magnetically enriched Siglec-F^+ eosinophils (blood, spleen, stomach, colon, small intestine, adipose tissue, lung and uterus, *Il5*-tg), total BM cells (*Il5*-tg) or $CD4^+$ cells (colon, B6J) were performed using the BD Rhapsody Single-Cell Analysis System (BD, Biosciences). Cells were pooled from three to five mice per sample. Tissue processing and enrichment procedures are described above. Each preparation was assessed by flow cytometry to determine eosinophil viability and was subjected to morphological examination after cytospin and staining. Eosinophils were labelled with sample tags (633793 BD Mouse Single-Cell Multiplexing Kit) according to the manufacturer's protocol. In brief, for each condition, 10^6 cells were resuspended in staining buffer (1% BSA, 1% EDTA in PBS) and incubated with the respective Sample Tag for 20 min at room temperature. Cells were then transferred to a 5-ml

polystyrene tube, washed twice with 2 ml staining buffer and centrifuged at 400*g* for 5 min. Samples were resuspended in 1 ml staining buffer for counting. Next, 10,000 or 20,000 cells from up to 4 barcoded samples were pooled for a total of 60,000 cells and the mixture was centrifuged at 400*g* for 5 min. The pellet was resuspended in 650 BD Sample Buffer supplemented with 1:1,000 SUPERase in (20 U μl^{-1}; AM2694 Thermo Fisher Scientific) and NxGen Rnase Inhibitor (40 U μl^{-1}; 30281-2 Lucigen). BD Rhapsody cartridges were super-loaded with 60,000 cells each. Single cells were isolated with the BD Rhapsody Express Single-Cell Analysis System according to the manufacturer's recommendations (BD Biosciences). cDNA libraries were prepared using the BD Rhapsody Whole Transcriptome Analysis Amplification Kit (633801 BD Biosciences) following the BD Rhapsody System mRNA Whole Transcriptome Analysis (WTA) and Sample Tag Library Preparation Protocol (BD Biosciences). The final libraries were quantified using a Qubit Fluorometer with the Qubit dsDNA HS Kit (Q32851 Thermo Fisher Scientific). Library size distribution was measured with the Agilent high-sensitivity D5000 assay on a TapeStation 4200 system (5067-5592 Agilent Technologies). Sequencing was performed in paired-end mode (2 × 75 cycles) on a NovaSeq 6000 with NovaSeq 6000 SP Reagent Kit chemistry.

Data pre-processing and normalization

After demultiplexing of bcl files with Bcl2fastq v.2.20.0.422 (Illumina) and quality control, paired-end scRNA-seq FASTQ files were processed on the Seven Bridges Genomics platform with default parameters. Downstream analysis was conducted in R v.4.1.0 with the package Seurat v.4.0.3 (ref. [43]). All Seurat objects (one for each of the multiplexed samples) were merged and subjected to the same quality filtering. Cells with fewer than 200 or more than 2,500 detected genes were excluded from the analysis. After log normalization, the count data were scaled regressing for mitochondrial reads, and principal component analysis (PCA) was performed based on the 2,000 most variable features. Clustering and UMAP visualization were performed on the merged dataset using 50 principal components and a resolution of 0.3 for the shared nearest neighbour clustering algorithm. The clusters were annotated manually on the basis of marker gene expression. Epithelial and mesenchymal contaminants, as well as immune-cell clusters

not belonging to the eosinophil lineage, were excluded from downstream analysis. A cluster high in mitochondrial genes was excluded as well. The eosinophil space was analysed by subsetting clusters expressing eosinophil markers. The subsetted dataset was subjected to normalization, scaling and PCA as above. Clustering and UMAP visualization was performed using 20 principal components and a resolution of 0.3 for the shared nearest neighbour clustering algorithm. For the lung, uterus and adipose tissue dataset, batch correction was performed with Harmony[44] and epithelial genes (marker genes of epithelial cluster with pct.2 < 0.05) derived from excessive-cell-free RNA were removed from the counts.

Differential gene expression analysis, gene-set enrichment and score computation

To extract cluster markers, FindAllMarkers was executed with logfc.threshold and min.pct cut-offs set to 0.25. Top-ranked genes (by log fold change; logFC) were extracted for illustration. For differential gene expression, FindMarkers was applied with logfc.threshold and min.pct set to 0. Genes were subsequently filtered on the basis of Bonferroni-adjusted $P < 0.05$. Scores were computed with the AddModuleScore function. Genes used for the scores and signatures were manually curated from Gene Ontology (GO) terms and literature, and are listed in Supplementary Table 3. Cell-cycle scoring was performed with the CellCycleScoring algorithm from Seurat, using cell-cycle-related genes[45]. For gene-set enrichment analysis (GSEA), differentially expressed genes were pre-ranked in decreasing order by the negative logarithm of their *P* value, multiplied for the sign of their average logFC (in R, '- log(p_val)*sign(avg_log2FC)'). GSEA was performed on this pre-ranked list using the R package FGSEA (https://github.com/ctlab/fgsea/) with default parameters and the GO Biological Process database, made accessible in R by the package msigdbr (https://github.com/cran/msigdbr). The results were filtered for significantly enriched gene sets (Bonferroni-adjusted $P < 0.05$).

Trajectory inference and trajectory alignment

Trajectory inference was performed with Monocle 2.3.6 (refs. [19,46]) in R v.3.6.3. After creating a Monocle object using 'negbinomial.size()'

distribution and lowerDetectionLimit = 0.5, the analysis was performed using Seurat's top 2,000 variable features as ordering genes. Dimensionality reduction was performed using the DDTree method. To visualize the eosinophil differentiation, cluster annotations were projected on the inferred trajectories. Trajectory alignment of the BM–blood–colon trajectories was performed by applying dynamic time warping as described previously[22,47]. The steady-state and *C. rodentium*-challenge trajectories were set as the reference and query, respectively. Differentially expressed genes were identifying by using a full model of 'y ~ pseudotime*treatment' and a reduced model of 'y ~ pseudotime'.

RNA velocity and cell fate probabilities

Loom files were generated with velocyto[48] and dynamical velocities were computed with scvelo[20]. Fate probabilities were computed with CellRank[21] and plotted as pie charts (partition-based graph abstraction, PAGA).

Analysis of pathway and regulon activity

Pathway activity was calculated across eosinophil subsets with PROGENy v.1.13.2 (ref. [49]) with default parameters. Gene-regulatory activity was interrogated by applying SCENIC 1.2.4 (ref. [28]) with default parameters. In brief, after expression matrix filtering (minCountsPerGene = 3*.01*ncol(exprMat), minSamples = ncol(exprMat)*.01), and computing correlation, GENIE3 was applied to infer potential transcription factor targets. Co-expression networks were then calculated, regulons were created and their activity was scored in cells. Regulon activities were visualized as cluster averages using the R package ComplexHeatmap (ref. [50]).

Integration of datasets

Challenge, DSS and B6J datasets were integrated using Seurat's anchoring-based integration method using the steady-state object as reference dataset (reference.reduction = "pca", dims = 1:50).

Prediction of cell–cell interaction with CellPhoneDB

Ligand–receptor interaction analysis was performed using the Python package CellPhoneDB (v.2.0.0, Python v.3.8.5) following instructions from the GitHub repository (https://github.com/Teichlab/cellphonedb). In brief, the annotated Seurat object of isolated lamina propria immune cells from DSS-treated B6J mice was used to test the expression of known ligand–receptor interactions from the public repository of CellPhoneDB. Gene symbols were first converted from mouse to human using the biomart R package (v.2.46.3). Mean values representing the average ligand and receptor expression of annotated clusters were calculated on the basis of the percentage of cells expressing the gene, and the gene-expression mean. To determine the significance of observed means, P values were calculated using a null distribution of means calculated for randomly permuted annotated cluster labels. An interaction was considered significant if $P \leq 0.05$. Significant ligand–receptor interaction pairs between eosinophils and $CD8^+$ T cells or $CD4^+$ T cells were extracted, gene symbols were converted from human to mouse and their mean values were plotted using the plot_cpdb function from the ktplots R package (v.1.1.14) (https://github.com/zktuong/ktplots).

Plotting and statistical analysis

Statistical analysis and visualization were performed using R version 3.6.3 or 4.1.0. Statistical significance tests were performed as described in each figure legend. Unless stated otherwise, all tests were significant with Bonferroni-adjusted $P < 0.05$. Plots were generated with the R package ggplot2 (ref. [51]).

Flow cytometry, cell sorting and counting

Staining

For surface staining, cells were stained in PBS at 4 °C for 30 min with the fixable viability dye eFluor 780 (1:1,000, 65-0865-14 eBioscience) and a combination of the following antibodies (1:200, all from BioLegend; unless stated otherwise): anti-mouse CD45 BV650 (30-F11, 103151), CD11b BV510 (M1/70, 101263), MHC-II AF700 (M5/114.15.2, 107622), Ly6G

Percp-Cy5.5 (1A8, 127616), CD4 PerCP (RM4-5, 100538), TCRβ PE-Cy7 (H57-597, 109222), TCRβ PE-Cy7 (H57-597, 109228), CD80 BV605 (1:100, 16-10A1, 104729), PD-L1 PE-Cy7 (1:100, 10F.9G2, 124314), CD31 PE (390, 102408), CD45.2 BV785 (1:50, 104, 109839), CD9 PE (MZ3, 124805), CD54 BV711 (YN1/1.7.4, 116143), CD63 PE (1:100, NVG-2, 143904), CD95 PE-Cy7(SA367H8, 152607), Siglec-E PE (M1304A01, 677104), SCA-1 AF488 (D7, 108116), SCA-1 AF700 (D7, 108142), C-kit BV605 (ACK2, 135121), CD11c APC-Cy7 (N418, 117323), CLEC12a PE (5D3, 143404), CD49d FITC (R1-2, 103605), CD16/32 FITC (S17012B, 101305), CD3e Percp-Cy5.5 (145-2C11, 100328), CD8a APC (53-6.7, 100712), NK1.1 Percp-Cy5.5 (PK136, 108727), B220 Percp-Cy5.5 (RA3-6B2, 103236), Ter119 Percp (TER-119, 116227), Gr1 Percp (RB6-8C5, 108427), CD34 AF647 (RAM34, 560230), Siglec-F BV421 (E50-2440, 552681 BD Biosciences), Siglec-F PE (E50-2440, 552126 BD Biosciences), CD125 PE (T21, 558488 BD Biosciences), CD275 (HK5.3, 50598582 eBioscience) and T1/ST2 FITC (1:100, DJ8, 101001F MD Bioproductos GmbH). For T cell intracellular cytokine staining, cells were incubated for 3 h 15 min in complete IMDM medium containing 0.1 μM phorbol 12-myristate 13-acetate (P-8139 Sigma) and 1 μM ionomycin (I-0634 Sigma) with 1:1,000 Brefeldin A (00-4506-51 eBioscience) and GolgiStop solutions (51-2092KZ BD Biosciences) in a humidified incubator with 5% CO_2 at 37 °C. After surface staining, cells were fixed and permeabilized with the Cytofix/Cytoperm Fixation/Permeabilization Solution kit (512090KZ BD Biosciences) according to the manufacturer's instructions. Cells were then stained for 50 min with anti-mouse IL-17A APC (TC11-18H10.1, 506916), IFNγ BV421 (XMG1.2, 505830) and TNF FITC (MP6-XT22, 506 304) all from Biolegend at 1:100. Fc block (anti-CD16/CD32, 101302 Affymetrix) was included to minimize nonspecific antibody binding. Total leukocyte counts were determined by adding countBright Absolute Counting Beads (C36950 Life Technologies) to each sample before analysis. Samples were acquired in a LSRII Fortessa or FACS AriaIII 5L (BD Biosciences). For high-dimensional spectral flow cytometry analysis, cells were acquired on Cytek Aurora 5L (Cytek Biosciences) following 50 min staining at 4 °C with the antibodies described in Supplementary Table 5. For the Click-iT Plus EdU Alexa Fluor 647 Flow Cytometry Assay Kit (C10419 Thermo Fisher Scientific), the staining protocol was followed according to the

manufacturer's instructions. BD FACSDiva Software (BD Biosciences) was used for data acquisition and cell sorting.

Data analysis and plotting

Flow cytometry data analysis was performed with FlowJo software (v.10.7.1 Becton Dickinson). Cell counts, relative cell frequencies or MFI were used to generate graphical plots in GraphPad Prism (v.9.1.1, GraphPad). High-dimensional flow cytometry data were compensated and exported with FlowJo software (v.10) and the resulting FCS files were uploaded into Rstudio (v.4.0.3 R software environment). UMAPs were generated on stochastically selected cells from each sample and FlowSOM metaclusterings were performed for all the exported events as described previously[52].

Statistical analysis

All statistical analyses were performed with GraphPad Prism (v.9.1.1, GraphPad). Two-tailed unpaired Student's *t*-tests were used for comparing two groups, and comparisons of more than two datasets were done using a one-way analysis of variance (ANOVA) with Tukey's post-test. Differences were considered statistically significant when $P < 0.05$.

Isolation and culture of mouse BM-derived eosinophils

To generate mouse BM-derived eosinophils (BM-Eos), BM cell suspensions were seeded at a density of 10^6 cells per ml in RPMI-1640 medium supplemented with 20% heat-inactivated FBS, 25 mM HEPES (H0887 Sigma), 100 U ml^{-1} penicillin–streptomycin (P0781 Sigma), 2 mM glutamine (25030-024 Gibco), 1× NEAA (11140-035 Gibco), and 1 mM sodium pyruvate (11360070 Gibco). Cells were cultured in a humidified incubator with 5% CO_2, 37 °C, and were supplemented with 100 ng ml^{-1} mouse SCF (250-03 PeproTech) and 100 ng ml^{-1} mouse FLT3-Ligand (250-31L PeproTech) from day 0 to day 4, followed by differentiation with 10 ng ml^{-1} mouse rec-IL-5 (215-15 PeproTech) until day 13, as described[53]. Half of the medium was replaced and the cell concentration was adjusted to 10^6

cells per ml every other day. On day 8, cells were collected and moved to new flasks to remove adherent contaminating cells. On day 13, the nonadherent cells were collected and washed with PBS. Eosinophils were sorted and purity was assessed by flow cytometry (higher than 95%).

In vitro conditioning with supernatant of cultured colonic explants and cytokines

Supernatant of cultured colonic explants (colon CM) was prepared by culturing mid-colon sections (around 0.3 cm) from 6–12-week-old female and male mice (B6J) in 300 µl complete RPMI medium in a humidified incubator with 5% CO_2, 24 h at 37 °C. Flow-cytometry-purified eosinophils were magnetically isolated from blood and spleen (*Il5*-tg) or differentiated from the BM (B6J) and were kept in complete RPMI medium with recombinant mouse IL-5 (10 ng ml^{-1}, PeproTech). Cells were seeded in round-bottom 96-well plates at a density of 2×10^5 cells per well (100 µl) and conditioned for 12 h at 37 °C with cell-free colon CM (1:10 or at the indicated doses) or the following cytokines: IL-22 (10 ng ml^{-1}, 210-22 PeproTech), IL-25 (10 ng ml^{-1}, 210-17E PeproTech), TNF (10 ng ml^{-1}, 315-01A PeproTech) and IL-33 (20 ng ml^{-1} or at the indicated doses, PeproTech). The NF-κB inhibitor BAY11-7082 (B5556, Sigma) was added at a concentration of 5 µM and anti-IL-33 neutralizing antibody (AF3626, Biotechne) at 30 ng ml^{-1}. To study granule mobilization, magnetically enriched splenic eosinophils (*Il5*-tg) were treated overnight with colon CM (1:10) and flow-cytometry-sorted A-Eos were conditioned with IFNγ (20 ng ml^{-1}, PeproTech) for 90 min.

C. rodentium ICC180 viability assay

Flow-cytometry-purified BM-Eos (B6J) or magnetically enriched colonic, splenic and blood eosinophils (*Il5*-tg) from 6–12-week-old female and male mice were used for the assay. BM-Eos were conditioned overnight with colon CM (1:10) at 37 °C. Eosinophils were washed with PBS and transferred to a white flat-bottom 96-well plate (Corning) in antibiotic-free RPMI-1640 medium supplemented with 10% FBS and mouse IL-5 (10 ng ml^{-1}, PeproTech). A total of 10^8 bioluminescent *C. rodentium* bacteria (at

exponential phase, optical density at 600 nm ($OD_{600\ nm}$) = 1–1.5) was added to each well and luminescence was measured after 60 min on an Infinite 200 PRO plate reader (TECAN).

T cell proliferation assay

Flow-cytometry-purified BM-Eos (B6J) or magnetically enriched splenic eosinophils (*Il5*-tg) or A-Eos and B-Eos sorted from the GI tract (*Il5*-tg) were isolated from 6–12-week-old female and male mice. BM-Eos or spleen-derived eosinophils were conditioned overnight with colon CM (1:10) or treated with recombinant mouse IFNγ (10 ng ml^{-1}, PeproTech) and/or IL-33 (20 ng ml^{-1}, PeproTech), as indicated. Naive $CD4^+$ T cells were isolated from the lymph nodes of 6–12-week-old female and male mice (B6J), enriched with the MojoSort Mouse CD4 Naïve T Cell Isolation Kit (480040 BioLegend) and purified by flow cytometry. T cells were labelled with the CellTrace CFSE Cell Proliferation Kit (C34554 Thermo Fisher Scientific) following the manufacturer's instructions. T cells were then activated by CD3/CD28 T-activator Dynabeads (11131D Gibco) and co-cultured with eosinophils at a 1:1 ratio (2×10^5 total) for 4 days at 37 °C in complete RPMI medium supplemented with 10 ng ml^{-1} recombinant mouse IL-5 (PeproTech) and 20 ng ml^{-1} IL-2 (402-ML R&D). CFSE dilution was assessed by flow cytometry.

Antigen presentation assay

BM-Eos were isolated from 6–8-week-old female and male mice (B6J) and purified by flow cytometry. Eosinophils were conditioned overnight with colon CM, where indicated. Cells were washed in PBS and loaded with 300 n ml^{-1} of ovalbumin (OVA) residues 257–264 (S7951 Sigma) or 323–339 (O1641 Sigma) for 6 h in complete RPMI medium supplemented with 10 ng ml^{-1} recombinant IL-5 (PeproTech). T cells were sorted by flow cytometry and labelled with CellTrace CFSE Cell Proliferation Kit (C34554 Thermo Fisher Scientific) following the manufacturer's instructions. OT-I $CD8^+$ and OT-II $CD4^+$ T cells were obtained from the lymph nodes of 8–12-week-old female and male mice (OT-I and OT-II, respectively). T cells were co-cultured with eosinophils at a 1:1 ratio (2×10^5 total) for 4 days at 37 °C in

complete RPMI medium supplemented with 10 ng ml^{-1} recombinant mouse IL-5 (PeproTech) and 20 ng ml^{-1} IL-2 (402-ML R&D). CFSE dilution was assessed by flow cytometry.

Quantitative PCR with reverse transcription (qRT–PCR)

The RNA from cultured BM-Eos (B6J) or A-Eos and B-Eos sorted from the small intestine (*Il5*-tg) was isolated using the Direct-zol RNA MicroPrep kit (R2062 Zymo Research), whereas the RNA from magnetically enriched colonic, splenic and blood eosinophils from 6–12-week-old female and male mice (*Il5*-tg) was isolated using the RNeasy Mini kit (74106 QIAGEN). Both isolations were performed according to the manufacturer's instructions, including the on-column DNase 1 digestion step. Complementary DNA synthesis was performed using Superscript III reverse transcription (18080-044 QIAGEN). Gene expression was measured on a CFX384 Touch Real-Time PCR system (Bio-Rad, Second Derivative Maximum method analysis with high-confidence algorithm) by TaqMan Gene Expression Assays (4331182 Applied Biosystems by Thermo Fisher Scientific): *Cxcl2* (Mm00436450_m1), *Hprt* (Mm03024075_m1), *Gapdh* (Mm99999915_g1), *Cd274* (Mm03048248_m1), *Cd80* (Mm00711660_m1), *Ahr* (Mm00478932_m1), *Nfkb1* (Mm00476361_m1), *Nfkb2* (Mm00479807_m1), *Rela* (Mm00501346_m1), *Tnfa* (Mm00443258_m1), *Il1b* (Mm00434228_m1) and *Ptgs2* (Mm00478374_m1). Gene-expression levels for each sample were normalized to *Hprt* or *Gapdh* expression. Mean relative gene expression was determined, and the differences calculated using the 2ΔC(t) method.

Bulk RNA sequencing

BM-Eos were isolated from seven-to-eight-week-old female and male mice (B6J), differentiated and purified by flow cytometry. Cells were plated at a density of 5×10^5 cells per well (250 μl) and conditioned overnight with recombinant IL-33 (20 ng ml^{-1} PeproTech) and/or IFNγ (15 ng ml^{-1} PeproTech). RNA isolation was performed with the RNeasy Mini kit (74106 QIAGEN) according to the manufacturer's instructions, including the on-column DNase 1 digestion step. RNA quality was assessed by Tapestation (Agilent). Library preparation was performed with the Illumina TruSeq

RNA Kit. RNA sequencing was performed on the Illumina Novaseq 6000 (200 Mio reads), single-end read 100 bp. Reads were quality-checked with FastQC. Read alignment to the reference genome Mus_musculus.GRCm39 and read count was performed on the Support Users for SHell script Integration (SUSHI) framework[54], with the RSEMApp application. Filtering and differential expression testing were performed with edgeR (ref. [55]). The package pheatmap (ref. [56]) was used to generate heat maps.

Immunofluorescence

Mouse colonic sections

The colon of 7–8-week-old female and male mice (B6J) was dissected out, flushed in PBS and fixed 3 h in PFA (4% in PBS) at 4 °C, followed by overnight incubation in sucrose (30% w/v in 4% PFA) at 4 °C. Tissue was embedded in Tissue-Tek OCT Compound (Sakura, 4583) and stored at −80 °C. Tissue from three or four mice was cryosectioned (8 µm) onto the same microscope slide, washed in PBS and incubated for 1 h in blocking solution (2.5% BSA, 5% heat-inactivated normal goat serum, 0.1% Tween-20 in PBS) at room temperature. Slides were incubated overnight in blocking solution with the following primary antibodies (1:100): rat anti-mouse Siglec-F (E50-2440, 552126 BD Biosciences), Armenian hamster anti-mouse CD80 (16-10A1, 104729 Biolegend) and rabbit anti-mouse pNF-κB p65 (Ser536) (93H1,3033S Cell Signalling). After washing three times with PBST (0.1% Tween in PBS), the following secondary antibodies were added (1:400 in blocking solution) to the slides for 1 h at room temperature: AlexaFluor goat anti-rat 594 (A-11007), AlexaFluor goat anti-hamster 647 (A-21451) and AlexaFluor goat anti-rabbit 488 (A-11008), all from Thermo Fisher Scientific. Slides were washed four times for 5 min with PBST, and DAPI (D9542 Sigma, 1:1,000) was added to the third washing step. Slides were mounted in Prolog Gold (P36930 Invitrogen) and imaged on a Nikon Ti2-E inverted microscope, equipped with CrestOptics X-Light v3 confocal disk unit, Lumencor Celesta lasers and Photometrics Kinetix camera.

Human tissue microarrays

The microarrays CO245 and CO246 were obtained from TissueArray.Com. Deparaffinized sections were subjected to antigen retrieval in 2.4 mM sodium citrate and 1.6 mM citric acid, pH 6, for 25 min in a steamer. Sections were washed with PBST and blocked for 1 h at room temperature in blocking buffer (5% BSA, 5% heat-inactivated normal goat serum in PBST). Slides were incubated overnight at 4 °C with the following primary antibodies (1:100, in blocking buffer): mouse anti-human MBP (BMK-13, anti-human MBP (BMK-13, MCA5751 Bio-Rad) and rabbit anti-human PD-L1 (E1L3N, 13684S Cell Signalling). After washing three times with PBST (0.1% Tween in PBS), the following secondary antibodies were added (1:400 in blocking solution) to the slides for 1 h at room temperature: AlexaFluor goat anti-rabbit 594 and AlexaFluor goat anti-mouse 647 (Thermo Fisher Scientific). DAPI staining, mounting and imaging were performed as above.

Cytospins

A total of 10^5 FACS-enriched spleen, blood and GI-tract-derived eosinophils (*Il5*-tg) from 7–8-week-old female and male mice were resuspended in 100 µl 5% FCS-supplemented RPMI medium and cytospun for 5 min at 50*g* into a funnel. Slides were air-dried for 30 min, fixed with ice-cold methanol for 5 min and then left to air dry overnight. Slides were washed, incubated for 1 h in blocking solution and stained overnight at 4 °C with mouse anti-EPX antibody (MM25-82.2.1 1:200, provided by E. A. Jacobsen), followed by 1-h incubation at room temperature with AlexaFluor goat anti-mouse 647. DAPI staining, mounting and imaging were performed as above. EPX staining intensity was quantified across the cell diameter in Fiji (MultiPlot) for 15 cells per condition.

Image analysis for quantification of the active-to-basal ratio of eosinophils

The cores used for quantification as well as patient data are available in Supplementary Table 2. Cores were chosen on the basis of the presence of colonic epithelium. ND files were imported in Imaris 9.6.0 and spot objects were created in the green (MBP) and red (PD-L1) channels separately (estimated *XY* diameter = 7 µm, estimated *Z* diameter = 4 um, quality filter >

6). To quantify the co-expression of PD-L1 and MBP, the distance of each spot in the green channel to the nearest spot in the red channel was computed. Green spots (eosinophils) with distance to red spots < 4 µm were considered as active eosinophils (co-expressing PD-L1). Green spots with distance to red spots > 4 µm were considered basal eosinophils. The active-to-basal ratio was then computed by dividing the number of active by the number of basal eosinophils in each core. For localization analysis, the active-to-basal ratio in colon crypts of human and mouse tissue was calculated in manually drawn ROIs comprising the lower (basal) or upper (luminal) thirds.

Histological assessment of colitis

Transversal mid-colon sections (0.5 cm) were fixed overnight in buffered 10% formalin solution, followed by paraffin embedding. Sections were stained with H&E. Histopathology of the colon was scored in a blinded manner considering four categories (each scored on a scale of 0–3): epithelial hyperplasia or damage and goblet cell depletion; leukocyte infiltration in the lamina propria; submucosal inflammation and oedema; area of tissue affected. The final score presented (0–12) represents the sums of all categories.

In vitro genome-wide CRISPR inhibition screen

A total of 1.3 billion BM stem cells (BMSCs) from 10–16-week-old female and male mice ($n = 27$, dCas9-KRAB) were isolated as described above. BMSCs were then split in two replicates and each lentivirally transduced with an independently amplified genome-wide CRISPR inhibition library[57] (Addgene 83987). Five days after transduction, BFP^{+} BMSCs were FACS-enriched and their culture medium was supplemented with recombinant IL-5 (10 ng ml^{-1}, PeproTech). After six days of IL-5-mediated differentiation, BM-Eos were conditioned with colon CM overnight (1:10). $PD\text{-}L1^{+}CD80^{+}$ eosinophils were sorted, the genomic DNA was extracted and sgRNAs were target amplified. Library size distribution was measured with the Agilent high-sensitivity D5000 assay on a TapeStation 4200 system (5067–5592 Agilent Technologies). Sequencing was performed in single-end mode (75 cycles) on Illumina NextSeq. Reads were trimmed with cutadapt (ref. [58]) and

aligned to the sgRNA references with Bowtie2 (ref. [59]). MAGeCK (ref. [60]) was used for guide counting and paired testing.

Western blotting

BM-Eos were isolated from 8–10-week-old female and male mice (B6J), differentiated and purified by flow cytometry. Cells were conditioned with colon CM (1:10) or rec-IL-33 (20 ng ml^{-1} PeproTech) for 45 min, then lysed in RIPA buffer (R0278 Sigma) supplemented with 2 mM sodium orthovanadate (J60191.AE Thermo Fisher Scientific), 15 mM sodium pyrophosphate (J62052.AK Thermo Fisher Scientific), 10 mM sodium fluoride (447351000 Thermo Fisher Scientific), and 1× complete protease inhibitor cocktail (11836153001 Roche). Protein concentrations were determined by BCA assay (23227 Pierce), and equal amounts were separated by SDS–PAGE using 10% acrylamide gels followed by transfer onto nitrocellulose membranes (88018 Thermo Fisher Scientific). Membranes were probed with antibodies against vinculin (42H89L44, 700062 Thermo Fisher Scientific), phospho-p38 MAPK (Thr180/Tyr182, MA5-15218 Thermo Fisher Scientific) and phospho-p65 (Ser536, 93H1, 3033 Cell Signalling Technology).

Enzyme-linked immunosorbent assay (ELISA)

Proteins were extracted from colon samples homogenized in 450 µl RIPA lysis buffer (Thermo Fisher Scientific) supplemented with Na_3Vo_4 (100 mM), NaF (10 mM) and protease inhibitor cocktail (cOmplete, Mini Protease Inhibitor Tablets, 11836153001 Roche). The supernatant was collected and centrifuged at maximum speed for 10 min at 4 °C. Protein concentration was quantified with the Pierce BCA Protein Assay Kit (23225, Thermo Fisher Scientific). Plasma was isolated from blood in BD Microtainer tubes (365968, BD). Plates were coated overnight and the mouse IL-33 ELISA kit (88-7333-88 Thermo Fisher Scientific) was used to quantify the colon and plasma levels of IL-33 according to the manufacturer’s instructions.

LEGENDplex bead-based immunoassay

Proteins were extracted as described above. Colon and plasma levels of IFNγ, IL-22 and TNF were quantified using LEGENDplex MU Th17 Panel (7-plex) according to the manufacturer's instructions.

Molecular cartography

Sample preparation

Fresh frozen colon samples from three patients with ulcerative colitis were sectioned onto coverslips and processed by Resolve Biosciences.

Segmentation

Cellpose (v. 2.0.4) (ref. [61]) was used to segment nuclei in the DAPI images with the pretrained nuclei model and flow_threshold 0.5, cellprob_threshold −0.2. The nuclear segments were then expanded by 10 pixel (1.38 µm) using the 'expand_labels' function in scikit-image and transcripts were subsequently assigned to the expanded segments. Segments with fewer than three molecules or three genes detected were removed from the analysis.

Segmentation-free approach

To circumvent issues of segmentation, we used a transcript-focused approach in which we used spatial clusters of specific marker genes to represent cell types and investigate co-localization. For this, distances between individual transcripts of *CD4*, *SIGLEC8*, *CD8A*, *CD19*, *FOXP3* and *FCN1* were computed using Euclidean distances of the 2D coordinates. Hierarchical clustering was then applied to the distance matrix with average linkage to prevent chaining and a tree cut at height of 5 µm (hclust in the stats R package). We then used a k-d-tree based nearest neighbour search to identify the clusters in the surrounding area of each other cluster in a pre-defined radius of 10 µm as implemented in the R function 'nn2' (RANN v.2.6.1, searchtype='radius') with a sufficiently large k ($k = 41$). This approach runs in O(M logM) time and avoids computation of a distance matrix for thousands of objects. Finally, a neighbourhood graph was constructed from the resulting adjacency matrix in which vertices (transcript

clusters) are connected by edges if they are no further apart than 10 µm. From this graph the number of edges between different cell types was computed and compared to an empirical null distribution that was derived from randomly permuting the labels of the vertices ($m = 1{,}000$). This approach takes tissue composition and spatial structure into account and allows the computation of P values as $P = (b + 1)/(m + 1)$, where b is the number of times the permutation produced a more extreme number of edges between two cell types than observed and m the total number of permutations[62]. This was done for each slide and possible cell–cell interaction to derive a score that represents the fraction of images in which a specific interaction was significant, with the sign representing interaction or avoidance; visualization was adopted from ref. [63].

Graphical illustrations

Schematics of experimental workflows were created using a licensed version of Biorender.com.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Single-cell and bulk RNA-seq data generated during this study have been deposited at the Gene Expression Omnibus under the accession number GSE182001.

Code availability

The code used in this study is available at https://github.com/Moors-Code/Eosinophils_scRNASeq.

References

1. Marichal, T., Mesnil, C. & Bureau, F. Homeostatic eosinophils: characteristics and functions. *Front. Med.* **4**, 101 (2017).

2. Blanchard, C., Wang, N. & Rothenberg, M. E. Eosinophilic esophagitis: pathogenesis, genetics, and therapy. *J. Allergy Clin. Immunol.* **118**, 1054–1059 (2006).

3. Humbles, A. A. et al. A critical role for eosinophils in allergic airways remodeling. *Science* **305**, 1776–1779 (2004).

4. Jenerowicz, D., Czarnecka-Operacz, M. & Silny, W. Peripheral blood eosinophilia in atopic dermatitis. *Acta Dermatovenerol. Alp Pannonica Adriat.* **16**, 47–52 (2007).

5. Raab, Y., Fredens, K., Gerdin, B. & Hällgren, R. Eosinophil activation in ulcerative colitis: studies on mucosal release and localization of eosinophil granule constituents. *Dig. Dis. Sci*. **43**, 1061–1070 (1998).

6. Chu, V. T. et al. Eosinophils promote generation and maintenance of immunoglobulin-A-expressing plasma cells and contribute to gut immune homeostasis. *Immunity* **40**, 582–593 (2014).

7. Jung, Y. et al. IL-1β in eosinophil-mediated small intestinal homeostasis and IgA production. *Mucosal Immunol.* **8**, 930–942 (2015).

8. Ignacio, A. et al. Small intestinal resident eosinophils maintain gut homeostasis following microbial colonization. *Immunity* **55**, 1250–1267 (2022).

9. Sugawara, R. et al. Small intestinal eosinophils regulate Th17 cells by producing IL-1 receptor antagonist. *J. Exp. Med.* **213**, 555–567 (2016).

10. Alhmoud, T. et al. Outcomes of inflammatory bowel disease in patients with eosinophil-predominant colonic inflammation. *BMJ Open Gastroenterol.* **7**, e000373 (2020).

11. Smillie, C. S. et al. Intra- and inter-cellular rewiring of the human colon during ulcerative colitis. *Cell* **178**, 714–730 (2019).

12. Sikkema, L. et al. An integrated cell atlas of the human lung in health and disease. Preprint at https://doi.org/10.1101/2022.03.10.483747 (2022).

13. Lee, N. A. et al. Expression of IL-5 in thymocytes/T cells leads to the development of a massive eosinophilia, extramedullary eosinophilopoiesis, and unique histopathologies. *J. Immunol.* **158**, 1332–1344 (1997).

14. Mahmudi-Azer, S., Downey, G. P. & Moqbel, R. Translocation of the tetraspanin CD63 in association with human eosinophil mediator release. *Blood* **99**, 4039–4047 (2002).

15. Khushman, M. et al. Exosomal markers (CD63 and CD9) expression and their prognostic significance using immunohistochemistry in patients with pancreatic ductal adenocarcinoma. *J. Gastrointest. Oncol.* **10**, 695–702 (2019).

16. Cohnen, A. et al. Surface CD107a/LAMP-1 protects natural killer cells from degranulation-associated damage. *Blood* **122**, 1411–1418 (2013).

17. Mesnil, C. et al. Lung-resident eosinophils represent a distinct regulatory eosinophil subset. *J. Clin. Invest.* **126**, 3279–3295 (2016).

18. Schwarzfischer, M. et al. TiO_2 nanoparticles abrogate the protective effect of the Crohn's disease-associated variation within the PTPN22 gene locus. *Gut* https://doi.org/10.1136/gutjnl-2021-325911 (2022).

19. Qiu, X. et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).

20. Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 (2020).

21. Lange, M. et al. CellRank for directed single-cell fate mapping. *Nat. Methods* **19**, 159–170 (2022).

22. McFaline-Figueroa, J. L. et al. A pooled single-cell genetic screen identifies regulatory checkpoints in the continuum of the epithelial-to-mesenchymal transition. *Nat. Genet.* **51**, 1389–1398 (2019).

23. Arnold, I. C. et al. Eosinophils suppress Th1 responses and restrict bacterially induced gastrointestinal inflammation. *J. Exp. Med.* **215**, 2055–2072 (2018).

24. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).

25. Masterson, J. C. et al. Eosinophil-mediated signalling attenuates inflammatory responses in experimental colitis. *Gut* **64**, 1236–1247 (2015).

26. Arnold, I. C. et al. The GM–CSF–IRF5 signaling axis in eosinophils promotes antitumor immunity through activation of type 1 T cell responses. *J. Exp. Med.* **217**, e20190706 (2020).

27. Griseri, T. et al. Granulocyte macrophage colony-stimulating factor-activated eosinophils promote interleukin-23 driven chronic colitis. *Immunity* **43**, 187–199 (2015).

28. Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).

29. Griesenauer, B. & Paczesny, S. The ST2/IL-33 axis in immune cells during inflammatory diseases. *Front. Immunol.* **8**, 475 (2017).

30. Kang, K. et al. Interferon-γ represses M2 gene expression in human macrophages by disassembling enhancers bound by the transcription factor MAF. *Immunity* **47**, 235–250 (2017).

31. Menzella, F. et al. Anti-IL5 therapies for severe eosinophilic asthma: literature review and practical insights. *J. Asthma Allergy* **13**, 301–313 (2020).

32. Adachi, O. et al. Targeted disruption of the *MyD88* gene results in loss of IL-1-and IL-18-mediated function. *Immunity* **9**, 143–150 (1998).

33. Hoshino, K. et al. Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the *Lps* gene product. *J. Immunol.* **162**, 3749–3752 (1999).

34. Rawlins, E. L., Clark, C. P., Xue, Y. & Hogan, B. L. M. The $Id2^+$ distal tip lung epithelium contains individual multipotent embryonic progenitor cells. *Development* **136**, 3741–3745 (2009).

35. Dent, L. A., Strath, M., Mellor, A. L. & Sanderson, C. J. Eosinophilia in transgenic mice expressing interleukin 5. *J. Exp. Med.* **172**, 1425–1431 (1990).

36. Lee, H.-M. et al. IFNγ signaling endows DCs with the capacity to control type I inflammation during parasitic infection through promoting $T\text{-}bet^+$ regulatory T cells. *PLoS Pathog.* **11**, e1004635 (2015).

37. Oboki, K. et al. IL-33 is a crucial amplifier of innate rather than acquired immunity. *Proc. Natl Acad. Sci. USA.* **107**, 18581–18586 (2010).

38. Townsend, M. J., Fallon, P. G., Matthews, D. J., Jolin, H. E. & McKenzie, A. N. T1/St2-deficient mice demonstrate the importance of T1/St2 in developing primary T helper cell type 2 responses. *J. Exp. Med.* **191**, 1069–1076 (2000).

39. Lee, J. J. et al. Defining a link with asthma in mice congenitally deficient in eosinophils. *Science* **305**, 1773–1776 (2004).

40. Doyle, A. D. et al. Homologous recombination into the eosinophil peroxidase locus generates a strain of mice expressing Cre recombinase exclusively in eosinophils. *J. Leukoc. Biol.* **94**, 17–24 (2013).

41. Diehl, G. E. et al. Microbiota restricts trafficking of bacteria to mesenteric lymph nodes by CX_3CR1^{hi} cells. *Nature* **494**, 116–120

(2013).

42. Wiles, S., Pickard, K. M., Peng, K., MacDonald, T. T. & Frankel, G. In vivo bioluminescence imaging of the murine pathogen *Citrobacter rodentium*. *Infect. Immun.* **74**, 5391–5396 (2006).

43. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).

44. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).

45. Kowalczyk, M. S. et al. Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. *Genome Res.* **25**, 1860–1872 (2015).

46. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).

47. Cacchiarelli, D. et al. Aligning single-cell developmental and reprogramming trajectories identifies molecular determinants of myogenic reprogramming outcome. *Cell Syst*. **7**, 258–268 (2018).

48. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).

49. Holland, C. H. et al. Robustness and applicability of transcription factor and pathway analysis tools on single-cell RNA-seq data. *Genome Biol.* **21**, 36 (2020).

50. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).

51. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).

52. Brummelman, J. et al. Development, application and computational analysis of high-dimensional fluorescent antibody panels for single-cell flow cytometry. *Nat. Protoc.* **14**, 1946–1969 (2019).

53. Dyer, K. D. et al. Functionally competent eosinophils differentiated ex vivo in high purity from normal mouse bone marrow. *J. Immunol.* **181**, 4004–4009 (2008).

54. Hatakeyama, M. et al. SUSHI: an exquisite recipe for fully documented, reproducible and reusable NGS data analysis. *BMC Bioinformatics* **17**, 228 (2016).

55. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

56. Kolde, R. pheatmap: Pretty Heatmaps. R version 1.0.12, https://cran.r-project.org/web/packages/pheatmap/index.html (2019).

57. Horlbeck, M. A. et al. Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *eLife* **5**, e19760 (2016).

58. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.J.* **17**, 10–12 (2011).

59. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

60. Li, W. et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014).

61. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* **18**, 100–106 (2021).

62. Phipson, B. & Smyth, G. K. Permutation P-values should never be zero: calculating exact P-values when permutations are randomly

drawn. *Stat. Appl. Genet. Mol. Biol.* **9**, Article39 (2010).

63. Lohoff, T. et al. Integration of spatial and single-cell transcriptomic data elucidates mouse organogenesis. *Nat. Biotechnol.* **40**, 74–85 (2022).

Acknowledgements

We thank S. Baghai Sain, D. Eletto, A. Ozga, F. Mhamedi Baccouche, A. Munitz and A. Müller for technical support and ideas; E. A. Jacobsen and the Mayo Clinic for providing PHIL and Eo-*Cre* mice and the anti-EPX antibody; and S. Tugues Solsona and C. Sparano for providing $Id2^{CreERT2};Rosa^{EYFP}$ mice. This study was supported by an SNSF Eccellenza Professorial Fellowship from the Swiss National Science Foundation (SNF) to I.C.A. (PCEFP3_187021) and A.E.M. (PCEFP3_ 181249). P.K. is supported by grants from the San Salvatore Foundation and the Helmut Horten Foundation. This study was further supported by the Helmsley Charitable Trust grant #1903-03791 to A.E.M. A.G. is supported by fellowship 2450 from the Hartmann-Müller Foundation and 34484 from the Olga Mayenfisch Foundation; I.G.-P. is supported by fellowship 20C197 from the Novartis Foundation for Medical-Biological Research; T.V. is supported by Czech Science Foundation grant 21-26025S and a National Institute for Cancer Research project (programme EXCELES, ID project no. LX22NPO5102), funded by the European Union (NextGenerationEU); T.V. and N.G.N. are fellows of the University of Zürich Research Priority Program (URPP) ‘Translational Cancer Research’.

Funding

Open access funding provided by Swiss Federal Institute of Technology Zurich.

Author information

Author notes

1. These authors contributed equally: Alessandra Gurtner, Costanza Borrelli

Authors and Affiliations

1. Institute of Experimental Immunology, University of Zürich, Zürich, Switzerland

   Alessandra Gurtner, Ignacio Gonzalez-Perez, Nicolás G. Núñez, Daniel Crepaz, Deeksha Raju, Burkhard Becher & Isabelle C. Arnold

2. Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland

   Costanza Borrelli, Karsten Bach, Ilhan E. Acar, Kristina Handler, Atefeh Lafzi & Andreas E. Moor

3. Institute of Pathology, University of Bern, Bern, Switzerland

   Vivian P. Vu & Philippe Krebs

4. Institute of Molecular Cancer Research, University of Zürich, Zürich, Switzerland

   Kristin Stirm

5. Institute of Physiology, University of Zürich, Zürich, Switzerland

   Julia Gschwend & Christoph Schneider

6. Department of Molecular Life Sciences, University of Zürich, Zürich, Switzerland

   Konrad Basler & Tomas Valenta

7. Institute for Food, Nutrition and Health, D-HEST, ETH Zürich, Zürich, Switzerland

   Emma Slack

8. Botnar Research Center for Child Health, Basel, Switzerland

   Emma Slack

9. Institute of Molecular Genetics of the Czech Academy of Sciences, Prague, Czech Republic

   Tomas Valenta

Contributions

A.G., C.B., I.G.-P., N.G.N., D.C., K.H., V.P.V., A.L., K.S., T.V., D.R. and J.G. performed experiments and analyses. C.B., N.G.N., K. Bach, I.E.A. and K.H. performed bioinformatics analyses. I.C.A., A.E.M., B.B., P.K., C.S., E.S. and K. Basler designed and supervised experiments. I.C.A., C.B. and A.G. wrote the manuscript.

Corresponding authors

Correspondence to Andreas E. Moor or Isabelle C. Arnold.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Dominic Grün, Ari Molofsky and Marc Rothenberg for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 scRNA-seq reveals five distinct eosinophil subpopulations.

a, Experimental workflow of scRNA-seq. **b**, UMAP of all sequenced single-cell transcriptomes passing quality control, clustered and annotated manually based on marker gene expression. **c**, Distribution of unique molecular identifiers (nUMI, log10 normalized), genes (nGenes, log10 normalized) and mitochondrial gene fraction (mitoRatio, log10 normalized) per cell across samples. **d**, Expression density of canonical eosinophil marker genes. **e**, Subset organ distribution. Dashed lines indicate eosinophil subsets from Fig. 1a. **f**, Significantly enriched (adjusted $P < 0.05$) GSEA terms across clusters. Kolmogorov–Smirnov test. **g**, Left: cell-cycle score. Middle: stemness score. Right: granulogenesis score. Data represents mean ± SD. Two-sided Wilcoxon test (n = 3, *Il5*-tg). **h**, Expression of cell-cycle genes across eosinophil subsets. Rows are genes and columns are single cells, coloured by scaled expression. **i**, Expression of *mKi67*, *Epx* and *S100a6* over pseudotime. **j**, Receptor gene expression in A-Eos and B-Eos. **k**, Immune-regulatory score across subsets. Data represents mean ± SD. Two-sided Wilcoxon test (n = 3, *Il5*-tg). Genes used for scores and signatures are listed in Supplementary Table 3.

Extended Data Fig. 2 PD-L1 and CD80 expression define active eosinophils in the GI tract.

a, GI surface marker gene expression in A-Eos and B-Eos. **b**, UMAP showing the normalized protein expression intensity of eosinophil surface markers (n = 4, B6J). **c**, Frequencies of A-Eos as assessed by flow cytometry (n = 4–6, B6J). Data represents mean ± SEM. One-way ANOVA. Data pooled from two independent experiments. **d**, Mean fluorescence intensity (MFI) of CD9, CD31, CD54 and CD95 across colonic eosinophil subsets. FMO: fluorescence minus one. **e**, Expression of A-Eos markers, normalized

to *Gapdh* in A-Eos and B-Eos sorted from the small intestine (n = 4, *Il5*-tg). Data represents mean ± SEM. Two-tailed unpaired Student's *t*-test. **f**, Frequencies of $CD63^{+}$, $CD9^{+}$ and $CD107a^{+}$ cells in A-Eos and B-Eos as assessed by flow cytometry (n = 6-7, B6J). Data represents mean ± SEM. Two-tailed unpaired Student's *t*-test. Data pooled from two independent experiments. **g**, EPX immunofluorescence in sorted blood and spleen eosinophils (n = 3, *Il5*-tg). Nuclei are stained with DAPI. Scale bar, 10 µm. **h**, Schematic representation of basal (lower) and luminal (upper) third of the mucosa. **i**, Representative immunofluorescence images of Siglec-F and CD80 in the mouse colon (n = 3, B6J). Arrows mark Siglec-$F^{+}CD80^{+}$ A-Eos (red) and Siglec-$F^{+}CD80^{−}$ B-Eos (green). Nuclei are stained with DAPI. Dashed lines delimit the border of luminal and basal third. Scale bar, 20 µm.

Extended Data Fig. 3 PD-$L1^{+}CD80^{+}$ A-Eos are specific to the mouse GI and enriched in human IBD.

a, Representative FACS plots of PD-$L1^{+}CD80^{+}$ and PD-$L1^{−}$ $CD80^{−}$ eosinophils (n = 3, B6J). Numbers indicate % of eosinophils. **b**,**c**, UMAP of eosinophil transcriptomes (shown in Fig. 1a) including those isolated from uterus, lung and adipose tissue (n = 4, *Il5*-tg). Cells coloured by organ (**b**) and by cluster (**c**). **d**, Subset distribution across organs (% of eosinophils). **e**, List of shared or unique markers (logFC > 0.5, *P* adjusted < 0.05) between A-Eos and tissue eosinophils. Non-parametric Wilcoxon rank sum test (FindMarkers function in Seurat). **f**, Representative FACS plots of PD-$L1^{+}CD80^{+}$and PD-$L1^{−}CD80^{−}$ eosinophils in HDM- or PBS-treated mice (n = 2, B6J). Numbers indicate % of eosinophils. **g**, Left: Gating strategy used to identify resident (rEos) and inflammatory (iEos) eosinophils as described by[17]. Right: quantification of PD-$L1^{+}CD80^{+}$ and PD-$L1^{−}CD80^{−}$ eosinophils in rEos and iEos. Medians are shown. **h**, MBP and PD-L1 immunofluorescence staining of human tissue microarrays. Representative cores from a healthy individual and a patient with Crohn's disease are shown (n = 5). Scale bars, 500 µm (core overview) and 10 µm (high magnification insets).

Extended Data Fig. 4 Challenge infection induces a compositional shift toward the A-Eos cluster.

a, Left: Representative FACS plots of the A-Eos and PD-L1^{-} CD80^{-} eosinophils. Numbers indicate % of eosinophils. Right: Absolute counts of A-Eos of mice shown in Fig. 2a–c. Medians are shown. Two-tailed unpaired Student's *t*-test. **b**, Data integration of challenge datasets (dark-red dots, n = 4, *Il5*-tg). Steady-state dataset (grey) used as a reference. Corresponding steady-state organs shown in black. **c**, Data integration of DSS dataset (dark-green dots, n = 3, B6J). Steady-state dataset (grey) is used as a reference. Sstate colon shown in black. **d**, UMAP of integrated (left) and merged (right) steady-state (grey) and challenge (red dots) datasets. **e**, Left: BM–blood-colon eosinophil Monocle trajectory at steady state and following *C. rodentium* infection. Each dot represents a single cell coloured by cluster identity. Right: RNA velocities (scvelo) in BM, blood and colon dataset as steady state and during *C. rodentium* infection. **f**, Significant DEGs (logFC > 0.5, adjusted $P < 0.05$) of circulating eosinophils found in the colon vs in the blood of *C. rodentium*-infected mice. Non-parametric Wilcoxon rank sum test (FindMarkers function in Seurat). **g**, Single-cell fate probabilities as calculated by CellRank and summarized for each cluster as a pie chart. Arrows represent velocity flow. Cells and pie charts coloured by cluster identity. **h**, Top: Workflow of *in vitro* conditioning. Bottom: A-Eos frequencies after conditioning with increasing doses of colon CM. Input: BM-derived (n = 5, B6J), blood (n = 5, *Il5*-tg) and splenic (n = 5, *Il5*-tg) eosinophils. Medians are shown. One-way ANOVA. **i**, Left: EYFP^{+} eosinophil frequencies over time across organs after single tamoxifen pulse in *Id2*^{CreERT2}*;Rosa*$^{\text{EYFP}}$ mice. Data represents mean ± SD. Right: Frequency of A-Eos and B-Eos in colonic EYFP^{+} eosinophils at day 2 and 4 post tamoxifen injection (n = 3, *Id2*^{CreERT2}*;Rosa*^{EYFP}). Medians are shown. Two-tailed unpaired Student's *t*-test. **j**, Antimicrobial and granulogenesis signature expression in A-Eos. **k**, Gene expression over common pseudotime at steady state (grey) and upon *C. rodentium* infection (dark red). Dots indicate single cells, coloured by organ: BM (blue), blood (yellow) and colon (red). **l**, Edu^{+}/Edu^{-} eosinophil ratio in the colon of *C. rodentium*-infected and control B6J (n = 5) and *Il5*-tg (n = 3) mice at day 4 post EdU injection. Data represent mean ± SEM. Two-tailed unpaired Student's *t*-test. **m**, Frequencies of eosinophil progenitors (gated as Live CD45^{+}CD11b^{+}IL5Ra^{+}Lin-Sca1^{-}CD34^{+}) in *C. rodentium*-infected (n = 17) and control (n = 9) B6J mice. Medians are shown. Data pooled from two independent experiments. Two-tailed unpaired Student's *t*-test. **n**, MFI of

CD63 in colonic A-Eos and B-Eos of *C. rodentium*-infected and control mice (n = 6, B6J). Medians are shown. Two-tailed unpaired Student's *t*-test. **o**, EPX immunofluorescence of sorted A-Eos of *C. rodentium*-infected and control mice (n = 5, *Il5*-tg). Nuclei are stained with DAPI. Insets show protrusions. Scale bar, 10 µm.

Extended Data Fig. 5 A-Eos interact with T cells.

a, Ligand–receptor interactions between eosinophils and $CD4^+$ T cells (left) or $CD8^+$ T cells (right) predicted by CellPhoneDB. Dot size and colour indicate interaction mean. **b**,**c**, Representative H&E-stained colonic sections (**b**) and colitis score (**c**) in B6J (n = 17) and PHIL (n = 13) mice assessed by histopathological examination; data are pooled from two independent experiments. Medians are shown. Two-tailed unpaired Student's *t*-test. Scale bars, 100 µm. **d**, Frequencies of IFNγ, IL-17 and TNF-expressing colonic $CD4^+$ T cells of DSS-treated B6J (n = 17) and PHIL (n = 13) mice. Medians are shown. Two-tailed unpaired Student's *t*-test. **e**, Left: CFSE dilution of T cells co-cultured with BM-derived eosinophils conditioned as indicated and loaded with ovalbumin (OVA) peptide. Right: Representative FACS plots of the CFSE dilution. Numbers indicate % of CFSE dilution (n = 3, B6J). Data represents mean ± SEM. Two-tailed unpaired Student's *t*-test. **f**, Left: A-Eos ($PD\text{-}L1^+CD80^+$) and B-Eos ($PD\text{-}L1^-CD80^-$) frequencies in stomach, colon and small intestine of B6J (n = 5) and *Il5*-tg (n = 5) mice. Medians are shown. Two-tailed unpaired Student's *t*-test. Right: Representative FACS plots. Numbers indicate % of eosinophils. **g**, UMAP of B6J colonic eosinophils (orange) at steady state (n = 6) and during *C. rodentium* infection (n = 5) integrated in the *Il5*-tg dataset (grey). *Il5*-tg colonic eosinophils at steady state and during *C. rodentium* infection in black. **h**, Antimicrobial signature, IFNγ-regulated gene signature and antigen processing and presentation via MHC-I in B6J colon and B6J colon + *C. rodentium*. Data represents mean ± SD. Two-sided Wilcoxon test (n = 3). **i**, A-Eos frequencies in *H. pylori*-infected (stomach, n = 5) and *C. rodentium*-infected (colon, n = 4–7) B6J and *Il5*-tg mice, relative to uninfected controls. Medians are shown. Two-tailed unpaired Student's *t*-test. **j**,**k** Absolute counts of A-Eos and B-Eos in colon and small intestine of B6J mice treated with anti-IL-5 (**j**, n = 5) or anti-CCR3 (**k**, n = 5) neutralizing antibodies and

the respective isotype control. Medians are shown. Two-tailed unpaired Student's *t*-test.

Extended Data Fig. 6 A-Eos are induced by NF-κB signalling.

a, Regulon activity across clusters. **b**, Representative regulons projected on UMAP plot. Cells are coloured by binary regulon activity. **c**, Pathway activity across clusters according to PROGENy analysis. **d**, Gene expression relative to *Hprt* measured by qRT–PCR of eosinophils sorted from the blood (n = 6), spleen (n = 6) and colon (n = 4) of *Il5*-tg mice. Data represents mean ± SEM. One-way ANOVA. **e**, Representative images of pNF-κB p65 immunofluorescence staining in colonic eosinophils (n = 3, B6J). Arrows mark A-Eos (Siglec-F^+ $CD80^+$, red) and B-Eos (Siglec-F^+ $CD80^-$, green). Nuclei are stained with DAPI. Scale bar, 20 µm. **f**, A-Eos frequencies upon conditioning of BM-Eos with colon CM and/or NF-κB inhibitor (n = 5, B6J). Data represents mean ± SD. One-way ANOVA. **g**, Representative FACS plots of colonic A-Eos (PD-$L1^+$$CD80^+$) and PD-$L1^-$$CD80^-$ (B-Eos), relative to Fig 3d. Numbers indicate % of eosinophils. **h**, Left: Representative FACS plots of colonic A-Eos (PD-$L1^+$$CD80^+$) and PD-$L1^-$$CD80^-$ (B-Eos), relative to Fig 3e. Numbers indicate % of eosinophils. Right: MFI of Siglec-F and % CD63 in colonic A-Eos and B-Eos shown in Fig. 3e. Medians are shown. Two-tailed unpaired Student's *t*-test. **i**, Colonic A-Eos and B-Eos frequencies at steady state in B6J (n = 5) $Tlr2^{-/-}$ (n = 3) and $Tlr4^{-/-}$ (n = 7) mice. Medians are shown. One-way ANOVA.

Extended Data Fig. 7 IL-33 induces the accumulation of A-Eos in the colon.

a, Experimental workflow of the CRISPR inhibition screen. **b**, Log10 negative score per gene, as calculated by MAGeCK. *Cd80* and *Cd274* evidenced in orange. Genes involved in TNF signalling pathway via NF-κB in red, and MAPK signalling pathway in darkred. **c**, IL-33 concentrations measured by ELISA in colon of DSS-treated mice (n = 12, B6J) and colon and blood of *C. rodentium*-infected mice (n = 7, B6J), compared to untreated controls (n = 7, B6J). Medians are shown. Two-tailed unpaired Student's *t*-test. **d**, IFNγ, TNF and IL-22 concentrations measured by LEGENDplex in colon (left) and blood (right) of *C. rod*entium-infected mice (n = 7, B6J),

compared to untreated controls (n = 7, B6J). Medians are shown. Two-tailed unpaired Student's *t*-test. **e**, A-Eos (PD-L1$^+$CD80$^+$) frequencies upon conditioning of BM-Eos with colon CM, IL-22, IL-25, TNF or IL-33 (n = 4, B6J). Data are pooled from two independent experiments. Medians are shown. One-way ANOVA. **f**, A-Eos frequencies after conditioning with increasing doses of IL-33. Input: BM-derived (n = 5, B6J), blood (n = 5, *Il5*-tg) and splenic (n = 5, *Il5*-tg) eosinophils. Medians are shown. One-way ANOVA. **g**, Western blot of phospho-p38 and phospho-p65 upon conditioning of BM-Eos with colon CM or IL-33 (n = 3, B6J). **h**, Gene expression normalized to *Hprt* measured by qRT–PCR of BM-Eos upon conditioning with IL-33 (n = 4, B6J). Data represents mean ± SEM. Two-tailed unpaired Student's *t*-test. **i**, ST2 expression in BM-Eos upon IL-33 treatment (n = 4, B6J). Data represents mean ± SEM. Two-tailed unpaired Student's *t*-test. **j**, ST2 expression in colonic A-Eos and B-Eos (n = 5, B6J). Data represents mean ± SEM. Two-tailed unpaired Student's *t*-test. **k**, ST2 expression across organs (n = 5, B6J). Data represents mean ± SEM. One-way ANOVA. **l**, A-Eos frequencies upon conditioning of WT (n = 2, pooled B6J) or ST2$^{-/-}$ (n = 2, pooled) BM-Eos with colon CM or IL-33. Technical replicates and mean ± SEM are shown. Two-tailed unpaired Student's *t*-test. **m**, Left: Representative FACS plots of A-Eos (PD-L1$^+$ CD80$^+$) and PD-L1$^-$CD80$^-$ eosinophils in the blood (top) and spleen (bottom). Numbers indicate % of eosinophils. Right: A-Eos frequencies in mice treated with IL-33 (n = 6-7, B6J). Medians are shown. Two-tailed unpaired Student's *t*-test. **n**, A-Eos and B-Eos frequencies in the indicated organs of B6J (n = 7) and *Il33*$^{-/-}$ (n = 5) mice at steady state. Medians are shown. Two-tailed unpaired Student's *t*-test.

Extended Data Fig. 8 A-Eos co-localize with CD4$^{\pm}$ T cells in human IBD.

a, Regulon activity in A-Eos across conditions (n = 4, *Il5*-tg). **b**, Multidimensional scaling (MDS) plot of bulk RNA-seq samples shown in Fig.4a. **c**, Heat map of signature gene expression across conditions of samples shown in Fig. 4a. **d**, A-Eos (PD-L1$^+$CD80$^+$) frequencies upon treatment of BM-Eos with IL-33 and/or IFNγ. (n = 4, B6J). Data represents mean ± SEM. One-way ANOVA. **e**, EPX immunofluorescence of A-Eos

upon exposure to IFNγ for 90 min. Splenic eosinophils were magnetically enriched (n = 2, *Il5*-tg), treated overnight with colon CM and A-Eos sorted by flow cytometry. Scale bar, 10 µm. **f**, Frequencies of PD-L1^{+} and CD80^{+} in colonic eosinophils of WT (n = 6, B6J) and Eo-*Cre;Ifngr$^{fl/fl}$* mice (n = 4) upon *C. rodentium* infection, relative to uninfected controls (n = 2, B6J). Medians are shown. Two-tailed unpaired Student's *t*-test. **g**, Left: UMAP of single-cell eosinophil transcriptomes isolated from the colon of anti-IFNγR-treated, *C. rodentium*-infected or control *Il5*-tg mice (n = 3). Middle: expression of IFNγ target genes. Right: Expression of granule and antimicrobial signatures. Data represents mean ± SD. Two-sided Wilcoxon test. **h**, Observed vs. expected number of contacts between clusters of *SIGLEC8* and *CD4* molecules shown per slide. P Values are computed based on a two-sided permutation test (see Methods). **i**, Proportions of segmented cells expressing *SIGLEC8* only (blue) or co-expressing both *SIGLEC8* and *CD4* (red) across slides. Dotted horizontal line shows mean. **j**, Mean count per slide of molecules of a given transcript in the proximity (<10 µm) of *SIGLEC8* RNA molecules spatially associated with *CD4* molecules vs *SIGLEC8* molecules not associated with *CD4* molecules. The central line in the box plot represents the median count per slide, the lower and upper hinge corresponds to the first quartiles and the whisker extends from the hinge to the smallest or largest value no further than 1.5 x IQR from the hinge. Two-sided paired Wilcoxon test (17 ROIs, n = 4 patients).

Supplementary information

Supplementary Information

Supplementary Discussion and Supplementary References.

Reporting Summary

Supplementary Figure

Uncropped western blot images.

Supplementary Figure

Gating strategy used for eosinophil flow cytometric analysis and sorting.

Supplementary Table 1

List of subset markers. List of top 100 markers for each eosinophil subset generated with FindAllMarkers function in Seurat.

Supplementary Table 2

Patient information of TMA samples. Patient gender, age and pathology information.

Supplementary Table 3

Genes for signature. List of genes used to define scores and signatures, with respective references.

Supplementary Table 4

List of DEGs of bulk RNA-seq. Differentially expressed genes per condition.

Supplementary Table 5

Antibody list. List of antibodies (antigen, clone, fluorochrome, dilution and manufacturer) used for high-dimensional spectral flow cytometric analysis.

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the

material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Gurtner, A., Borrelli, C., Gonzalez-Perez, I. *et al.* Active eosinophils regulate host defence and immune responses in colitis. *Nature* **615**, 151–157 (2023). https://doi.org/10.1038/s41586-022-05628-7

- Received: 18 March 2022
- Accepted: 06 December 2022
- Published: 12 December 2022
- Issue Date: 02 March 2023
- DOI: https://doi.org/10.1038/s41586-022-05628-7

Share this article

Anyone you share the following link with will be able to read this content:

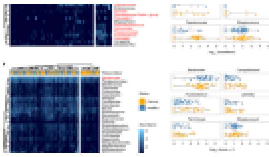
Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

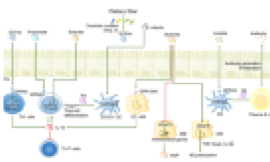
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-022-05628-7



The gut microbial metabolite formate exacerbates colorectal cancer progression

18 April 2022

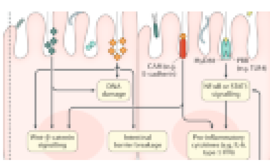
Dominik Ternes, Mina Tsenkova, … Elisabeth Letellier



Demystifying the manipulation of host immunity, metabolism, and extraintestinal tumors by the gut microbiome

12 October 2019

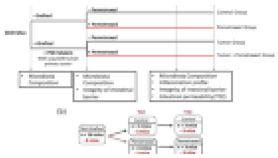
Ziying Zhang, Haosheng Tang, … Yongguang Tao



Gut microbiota in colorectal cancer: mechanisms of action and clinical applications

25 September 2019

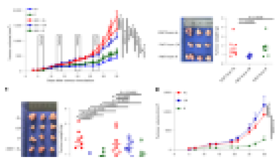
Sunny H. Wong & Jun Yu



Impact of pemetrexed chemotherapy on the gut microbiota and intestinal inflammation of patient-lung-derived tumor xenograft (PDX) mouse models

04 June 2020

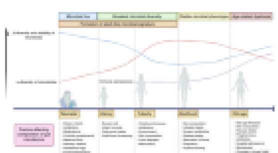
Cindy Pensec, Florence Gillaizeau, … Françoise Le Vacon



The antitumour effects of caloric restriction are mediated by the gut microbiome

16 January 2023

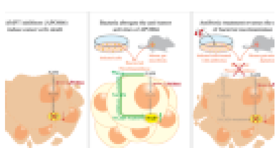
Yu-Qin Mao, Jia-Ting Huang, … Li-Shun Wang



The gut microbiome: what the oncologist ought to know

14 July 2021

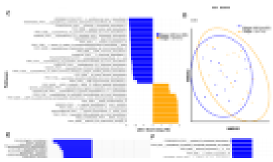
K. A. Lee, M. K. Luong, … T. D. Spector



[Gut microbiota severely hampers the efficacy of NAD-lowering therapy in leukemia]

08 April 2022

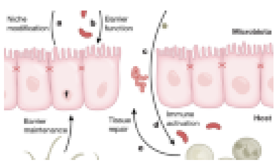
Oussama ElMokh, Saki Matsumoto, … Michel A. Duchosal



[Intestinal microbiota influences clinical outcome and side effects of early breast cancer treatment]

07 May 2021

Safae Terrisse, Lisa Derosa, … Laurence Zitvogel



[Host–microbiota maladaptation in colorectal cancer]

23 September 2020

Alina Janney, Fiona Powrie & Elizabeth H. Mann

- Article
- Open Access
- Published: 22 February 2023

Microbiota-derived 3-IAA influences chemotherapy efficacy in

pancreatic cancer

- Joseph Tintelnot ,
- Yang Xu,
- Till R. Lesker,
- Martin Schönlein,
- Leonie Konczalla,
- Anastasios D. Giannou,
- Penelope Pelczar,
- Dominik Kylies,
- Victor G. Puelles,
- Agata A. Bielecka,
- Manuela Peschka,
- Filippo Cortesi,
- Kristoffer Riecken,
- Maximilian Jung,
- Lena Amend,
- Tobias S. Bröring,
- Marija Trajkovic-Arsic,
- Jens T. Siveke,
- Thomas Renné,
- Danmei Zhang,
- Stefan Boeck,
- Till Strowig,
- Faik G. Uzunoglu,
- Cenap Güngör,
- Alexander Stein,
- Jakob R. Izbicki,
- Carsten Bokemeyer,
- Marianne Sinn,
- Alec C. Kimmelman,
- Samuel Huber &
- …
- Nicola Gagliani

Show authors

Nature **volume 615**, pages 168–174 (2023)

- 18k Accesses
- 1 Citations
- 184 Altmetric
- Metrics details

Abstract

Pancreatic ductal adenocarcinoma (PDAC) is expected to be the second most deadly cancer by 2040, owing to the high incidence of metastatic disease and limited responses to treatment[1,2]. Less than half of all patients respond to the primary treatment for PDAC, chemotherapy[3,4], and genetic alterations alone cannot explain this[5]. Diet is an environmental factor that can influence the response to therapies, but its role in PDAC is unclear. Here, using shotgun metagenomic sequencing and metabolomic screening, we show that the microbiota-derived tryptophan metabolite indole-3-acetic acid (3-IAA) is enriched in patients who respond to treatment. Faecal microbiota transplantation, short-term dietary manipulation of tryptophan and oral 3-IAA administration increase the efficacy of chemotherapy in humanized gnotobiotic mouse models of PDAC. Using a combination of loss- and gain-of-function experiments, we show that the efficacy of 3-IAA and chemotherapy is licensed by neutrophil-derived myeloperoxidase. Myeloperoxidase oxidizes 3-IAA, which in combination with chemotherapy induces a downregulation of the reactive oxygen species (ROS)-degrading enzymes glutathione peroxidase 3 and glutathione peroxidase 7. All of this results in the accumulation of ROS and the downregulation of autophagy in cancer cells, which compromises their metabolic fitness and, ultimately, their proliferation. In humans, we observed a significant correlation between the levels of 3-IAA and the efficacy of therapy in two independent PDAC cohorts. In summary, we identify a microbiota-derived metabolite that has clinical implications in the treatment of PDAC, and provide a motivation for considering nutritional interventions during the treatment of patients with cancer.

Main

Polychemotherapy, either with 5-fluorouracil (5-FU), irinotecan and oxaliplatin in combination with folinic acid (FOLFIRINOX), or with gemcitabine and nab-paclitaxel (GnP), is considered the standard of care for patients suffering from metastatic PDAC (mPDAC)[3,4]. However, less than half of all patients are responsive to the therapy, and patients who do not respond (NR patients) suffer from pain and eventually die within weeks[3]. Genetic alterations in PDAC poorly explain the differences between patients who respond to therapy (responder (R) patients) and NR patients[5,6], which leaves environmental factors—including the intestinal microbiota—as the potential mediators of chemotherapy efficacy. There is, therefore, an urgent need to identify environmental factors that might explain the differences between R and NR patients so that new concepts can be developed for future therapies.

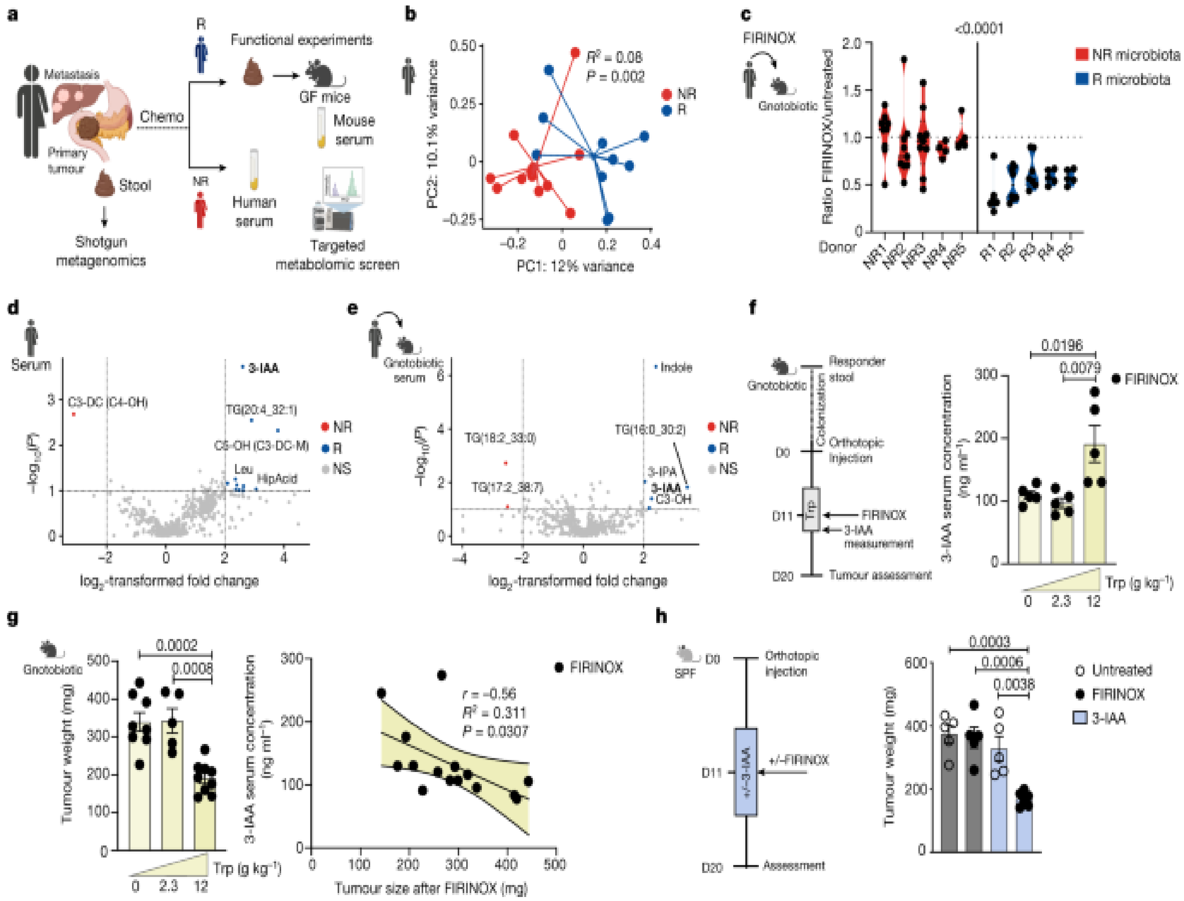
The intestinal microbiota has been shown to induce a response to immunotherapy in patients with melanoma, and can be modulated by dietary habits[7,8,9,10]. In rare patients with localized PDAC who are long-term survivors, bacteria can translocate from the intestine into the tumour and control anti-tumour immune activation[11,12]. However, most patients suffering from aggressive immunotherapy-resistant mPDAC are treated with polychemotherapy, and it is at present unclear whether and how the microbiota or dietary habits affect its efficacy[1,13].

3-IAA induces a response to chemotherapy

We recruited 30 patients with mPDAC, of whom 23 did not receive antibiotics and provided enough sample material to allow the intestinal microbiota to be analysed before the beginning of chemotherapy treatment (Fig. 1a and Extended Data Fig. 1a). The cohort of patients was separated into R and NR patients mainly on the basis of radiological response or, in cases in which the disease stabilized or computed tomography (CT) scans were missing, on the basis of progression-free survival (PFS) and a decrease of serum tumour markers (see Methods for detailed criteria). R patients (n = 11) had a mean PFS of 40.9 weeks, which was significantly higher than the

PFS of 12.8 weeks for NR patients ($n = 12$), and the overall survival was 51.9 weeks for R and 26.4 weeks for NR patients (Extended Data Fig. 1b,c). The microbiota of R patients was distinct from the microbiota of NR patients (Fig. 1b and Extended Data Fig. 1d–f). To study a potential cause–effect relationship between the microbiota and the response to chemotherapy, we transferred microbiota from the first ten recruited R and NR patients into gnotobiotic mice, followed by orthotopic injection of *Pdx1*-Cre, LSL-*KRAS*^{G12D}, LSL-*Trp53*$^{R172H/+}$ (KPC) pancreatic cancer cells (Extended Data Fig. 2a–d). Notably, four patients were treated with FOLFIRINOX and one patient was treated with GnP in both groups (R and NR). Irrespective of the original donor treatment, we observed smaller tumours in mice that were colonized with the microbiota of R patients, but not in mice colonized with the microbiota of NR patients, after chemotherapy treatment (5-FU, irinotecan and oxaliplatin; FIRINOX) (Fig. 1c and Extended Data Fig. 2e). Given the potential translocation of intestinal bacteria into PDAC tumours[11], we next analysed the intratumoural bacteria by 16S rRNA sequencing. We could only detect intratumoural bacteria in 2 out of 12 tumours (Extended Data Fig. 2f), and thus hypothesized that the response to chemotherapy is indirectly controlled through circulating microbiota-derived metabolites.

Fig. 1: 3-IAA induces a response to FIRINOX in mouse models of PDAC.



a, The intestinal microbiota of 23 patients with mPDAC was sequenced before the start of chemotherapy (chemo). GF, germ-free. **b**, Principal coordinate analysis (PCoA), with Bray–Curtis dissimilarity matrix, of pre-treatment NR ($n = 12$) and R ($n = 10$; one patient was excluded after the quality control) microbiota. **c**, Gnotobiotic mice were colonized with five different R or NR microbiota, KPC cancer cells were orthotopically injected and mice were treated with FIRINOX or left untreated. Tumour weight is shown relative to the mean tumour weight of the untreated group of each experiment at day 20 after tumour cell injection ($n = 9$ NR1–3; $n = 8$ R2; $n = 7$ R1 and R3; $n = 5$ NR5, R4 and R5; $n = 4$ NR4; pooled from nine independent experiments). **d**,**e**, Volcano plots showing differentially abundant metabolites in the serum of three R and NR patients (**d**) and in the serum of gnotobiotic mice colonized with R or NR microbiota ($n = 3$ each, biological replicates) (**e**). NS, not significant. **f**, R-microbiota-colonized gnotobiotic mice were fed with the indicated concentration of tryptophan, and the 3-IAA serum concentration at the fourth day of dietary intervention is shown ($n = 5$ each). **g**, Left, tumour weight of orthotopic KPC tumours (n

= 8, 5 or 9) after FIRINOX treatment. Right, correlation between 3-IAA serum concentration and tumour weight of five randomly selected mice per dietary group from **f** (*n* = 15). **h**, SPF mice were orthotopically injected with KPC cells and treated with or without (+/−) 3-IAA and/or with or without FIRINOX, and tumour weight was assessed at day 20 of the experiment (*n* = 5 or 6). Each symbol represents one mouse. One out of two (**f**,**g**) or three (**h**) independent experiments is shown. Error bars indicate s.e.m. Significant *P* values are indicated and were determined by MANOVA (**b**), two-tailed nested *t*-test (**c**), fold change analysis and two-tailed *t*-test (**d**,**e**), simple linear regression and Pearson's *r* (**g**) and one-way ANOVA with Tukey's post-hoc test (**f**–**h**).

Source data

To address this, we analysed serum from R and NR patients and matching colonized gnotobiotic mice and performed a targeted metabolomic screen using liquid chromatography coupled to mass spectrometry. We found that the tryptophan metabolite 3-IAA was the most significantly enriched metabolite in R compared to NR patients (Fig. 1d). In line with this, 3-IAA was also enriched in the serum of gnotobiotic mice that were colonized with R compared to NR microbiota (Fig. 1e and Extended Data Fig. 2g).

To characterize which gut bacteria in R patients could contribute to increased production of 3-IAA, we analysed the abundance of common 3-IAA-producing bacterial strains[14] in the microbiota of R compared to NR patients. Out of the fifteen analysed 3-IAA producers, we found that *Bacteroides fragilis* and *Bacteroides thetaiotaomicron* were increased in R patients and confirmed their capacity to produce 3-IAA in vitro (Extended Data Fig. 2h,i).

We next wondered whether we could alter the serum levels of 3-IAA in mice and thereby affect the efficacy of chemotherapy by modulating the dietary concentration of tryptophan, the precursor of 3-IAA (ref. [15]). Because tryptophan can impair the development of anti-tumour immune responses and thus promote tumour growth[16,17], we first tested different lengths of dietary intervention. Fourteen days of tryptophan exposure promoted tumour growth, whereas a treatment period of only four days was not sufficient (Extended Data Fig. 2j,k). We therefore decided to choose a four-to-five-day

intervention for the rest of our experiments to avoid the pro-tumorigenic effect[16,17].

Second, we observed that four days of a high-tryptophan diet was already sufficient to increase the concentration of 3-IAA in the serum of gnotobiotic mice that were colonized with R microbiota (Fig. 1f). Moreover, when this dietary intervention was combined with FIRINOX, we observed a decreased tumour weight (Fig. 1g). In addition, 3-IAA serum levels correlated inversely with tumour weight (Fig. 1g). Notably, these tryptophan-mediated effects—that is, increased concentration of 3-IAA and decreased tumour weight—were lost in NR-microbiota-colonized mice (Extended Data Fig. 2l,m). These data suggest that the effect of the tryptophan-high diet is mediated by 3-IAA, but we cannot rule out additional mechanisms contributing to the observed increased response in this set of experiments. Therefore, we directly supplemented 3-IAA in specific-pathogen-free (SPF) mice and found that our intervention was sufficient to reach the concentrations measured in mice that were colonized with R microbiota, and to increase the efficacy of chemotherapy not only in SPF but also in NR-colonized gnotobiotic mice (Fig. 1h and Extended Data Fig. 3a–c). We also tested other microbiota-modulated metabolites that have implications in different gastrointestinal cancers, such as the secondary bile acid deoxycholic acid (DCA) and the primary bile acid glycocholic acid (GCA)[18,19]. In addition, we chose to test two other metabolites that were increased in R patients or R-microbiota-colonized gnotobiotic mice—namely, the indole derivate indole-3-propionic acid (IPA) and hippuric acid—in combination with FIRINOX. None of these metabolites led to a similar efficacy to that of 3-IAA (Extended Data Fig. 3d,e). Together, these data indicate that the microbiota-derived tryptophan metabolite 3-IAA is increased in the serum of humans and mice that respond to chemotherapy, and that the levels of 3-IAA can be modulated by dietary interventions with tryptophan. Moreover, 3-IAA rendered even chemotherapy-resistant PDAC susceptible to the treatment.

The effect of 3-IAA is licensed by myeloperoxidase

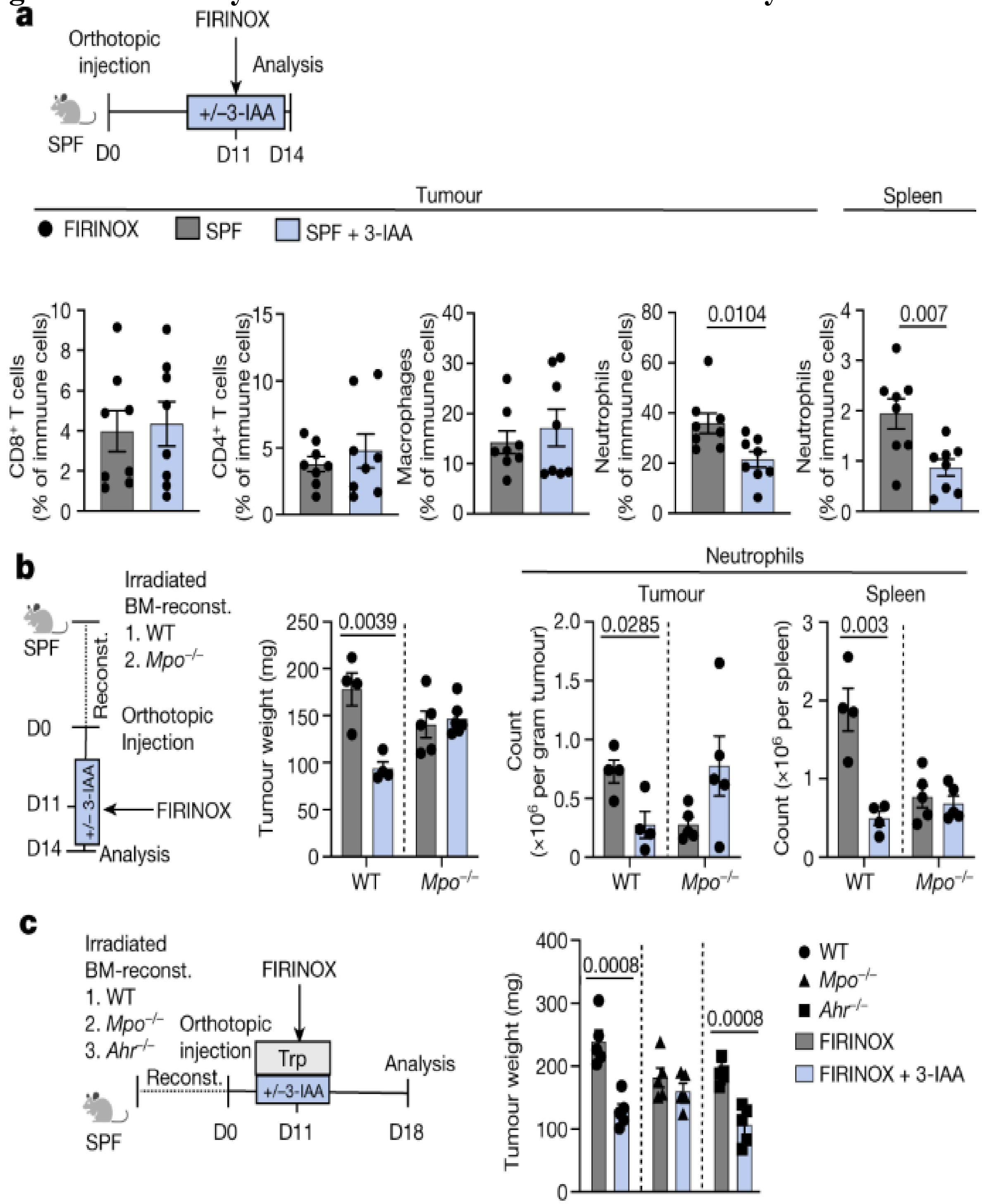
Microbiota-derived metabolites, especially tryptophan metabolites, have a crucial role in shaping innate and adaptive immunity[17], which in turn have

an important role in determining the efficacy of chemotherapy and prognosis in PDAC[11,20]. Therefore, we decided to analyse tumour-infiltrating immune cells in chemotherapy-naive or treated R- or NR-microbiota-colonized gnotobiotic mice. We consistently observed increased frequencies of $CD8^{+}$ T cells and decreased neutrophils in R- compared to NR-microbiota-colonized mice after chemotherapy. We did not observe any changes in tumour-infiltrating immune cells when comparing untreated mice (Extended Data Fig. 4a,b). $CD8^{+}$ T cells can induce the regression of PDAC when activated by the intestinal microbiota[11,21]. However, depleting $CD8^{+}$ or both $CD4^{+}$ and $CD8^{+}$ T cells through antibody treatments did not decrease the efficacy of chemotherapy in R-microbiota-colonized mice that received either a standard or a high-tryptophan diet (Extended Data Fig. 5a–d). Similarly, depletion of $CD4^{+}$ and $CD8^{+}$ T cells did not lower the efficacy of 3-IAA and FIRINOX in SPF mice (Extended Data Fig. 5e). Thus, the effect of 3-IAA does not seem to depend on the presence of T cells.

Neutrophils are highly abundant in PDAC, and a low neutrophil count (neutropenia) is associated with a good prognosis in mPDAC[22]. 3-IAA is specifically toxic to cells with a high concentration of myeloperoxidase (MPO), which is a hallmark of neutrophils[23]. Mechanistically, MPO can oxidize 3-IAA, inducing toxic products such as 3-methylene-2-oxindole (MOI)[24,25]. In line with this, we found that culturing bone-marrow-derived neutrophils, but not T cells (neutrophils have higher levels of MPO than T cells), with 3-IAA and FIRINOX led to reduced cell survival (Extended Data Fig. 6a). Moreover, adding MPO to bone-marrow-derived neutrophil cultures or using immature neutrophils (pre-neutrophils) that have higher levels of endogenous MPO (ref. [26] and Extended Data Fig. 6b) increased the toxicity of 3-IAA and FIRINOX (Extended Data Fig. 6c,d). By contrast, IPA did not increase the efficacy of FIRINOX in inducing neutrophil cell death in vitro (Extended Data Fig. 6e). Further characterization of the response of neutrophils to 3-IAA and FIRINOX revealed that the treatment induced neutrophil degranulation, which was measured as the release of MPO, and necrosis, but did not lead to the formation of neutrophil extracellular traps (NETs) or apoptosis (Extended Data Fig. 6f–h). Next, we wondered whether these results were also reproducible in vivo. The combination of 3-IAA and FIRINOX reduced the frequency and number of neutrophils in the tumour

and spleen of SPF mice (Fig. 2a and Extended Data Fig. 6i). Notably, this was not the case for 3-IAA treatment alone (Extended Data Fig. 6j).

Fig. 2: The efficacy of 3-IAA and FIRINOX is licensed by MPO.



a, SPF mice were orthotopically injected with KPC cells and treated with FIRINOX with or without 3-IAA, and tumours were analysed at day three after FIRINOX treatment ($n = 8$ each). Immune subsets of orthotopic tumours or respective spleens were determined by flow cytometry. $CD8^+$ T cells ($CD3^+CD8^+$); $CD4^+$ T cells ($CD3^+CD4$); macrophages ($CD11b^+F4/80^+$); neutrophils ($CD11b^+Ly6G^+$) are shown as relative to total living immune cells ($CD45^+$). **b**, Tumour weight and counts of immune subsets of orthotopic KPC tumours or respective spleens of irradiated and wild-type (WT; $n = 4$) or *Mpo*$^{-/-}$ ($n = 5$ or 6) bone-marrow (BM)-reconstituted mice treated and analysed as in **a**. **c**, Irradiated and wild-type, *Mpo*$^{-/-}$ or *Ahr*$^{-/-}$ bone-marrow-reconstituted mice received KPC cells orthotopically and were treated with FIRINOX or FIRINOX + 3-IAA for five days ($n = 5$ each). All mice received a four-day dietary intervention with a high-tryptophan diet. Tumour weight was assessed at day seven after FIRINOX treatment. Each symbol represents one mouse. Two independent experiments were pooled (**a**) or one out of two independent experiments (**b**,**c**) is shown. Error bars indicate s.e.m. Significant P values are indicated and were determined by two-tailed Mann–Whitney test (**a**) or two-tailed t-test (**b**,**c**).

Source data

Given the crucial role of neutrophil-derived MPO in the release of toxic 3-IAA products, we wondered whether this could explain the increased efficacy of 3-IAA and FIRINOX on tumour growth. To address this question, we established bone marrow chimeras from either wild-type or *Mpo*$^{-/-}$ mice. As expected, treatment with FIRINOX and 3-IAA resulted in the depletion of neutrophils and showed synergistic efficacy in wild-type reconstituted mice (Fig. 2b). However, in mice that were reconstituted with *Mpo*$^{-/-}$ bone marrow, the addition of 3-IAA did not result in smaller tumours (Fig. 2b), suggesting that MPO is essential for the efficacy of 3-IAA and FIRINOX. Further supporting the role of MPO, the 3-IAA oxidation product MOI was increased in tumours from wild-type SPF mice after treatment with 3-IAA and FIRINOX compared to that with FIRINOX alone (Extended Data Fig. 7a). In addition, the levels of MOI were strongly reduced in tumours from *Mpo*$^{-/-}$ bone-marrow-reconstituted mice (Extended Data Fig. 7b).

Finally, considering the key role of the aryl hydrocarbon receptor (AhR*)* as an intracellular receptor for 3-IAA and other indols[27], we wondered whether *Ahr*$^{-/-}$ bone-marrow-reconstituted mice are resistant to 3-IAA and FIRINOX. The size of the tumours in *Ahr*$^{-/-}$ bone-marrow-reconstituted mice was reduced to the same extent as that of tumours from wild-type reconstituted mice after treatment with 3-IAA and FIRINOX (Fig. 2c), arguing against a role of the AhR in mediating the efficacy of 3-IAA and FIRINOX treatment. Moreover, the AhR was also dispensable in cancer cells as *Ahr*-knockdown KPC cells responded to 3-IAA and FIRINOX treatment in vivo (Extended Data Fig. 7c).

These data show that the efficacy of 3-IAA and FIRINOX is licensed by immune cell-derived MPO, but not by AhR signalling.

The effect of 3-IAA depends on ROS and autophagy

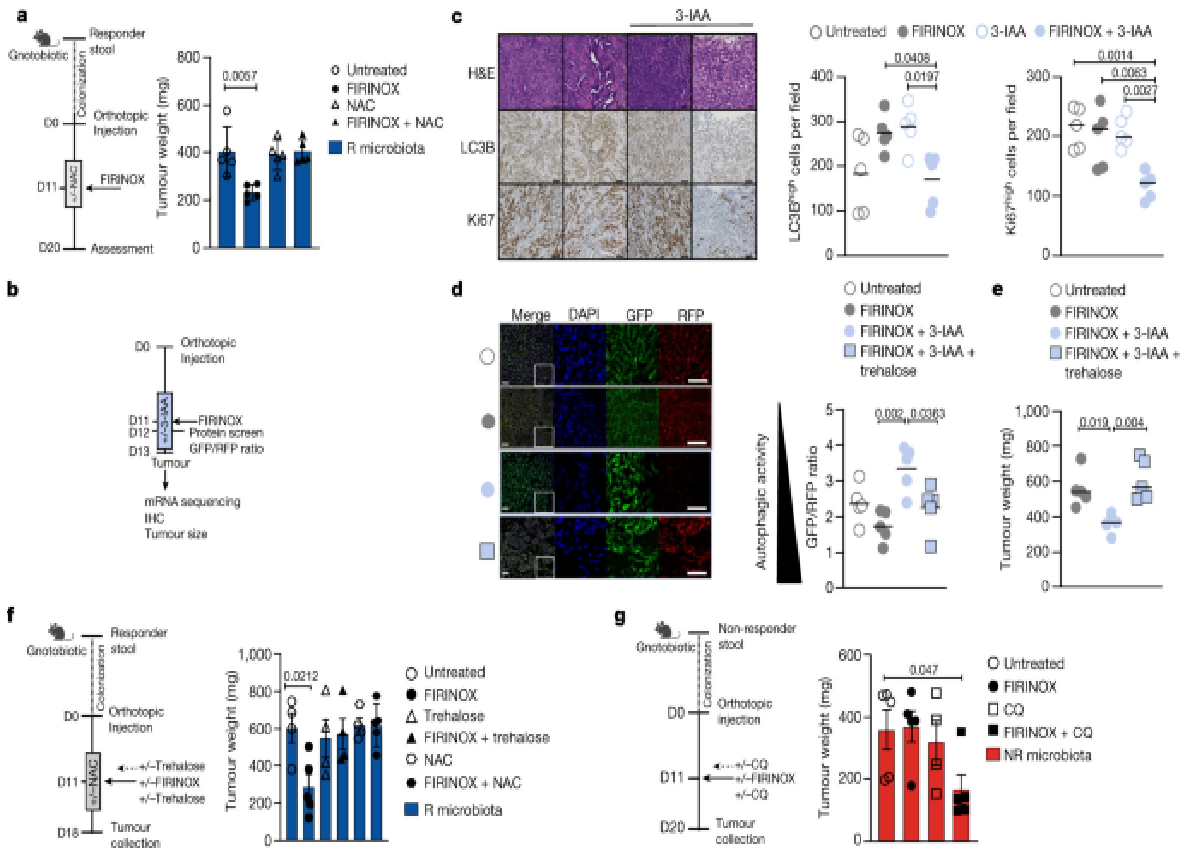
Oxidation of 3-IAA via MPO induces ROS in cultured neutrophils[23] and ROS are major mediators of chemotherapy-induced cell death[28]. Therefore, we hypothesized that the efficacy of 3-IAA and FIRINOX in the presence of MPO is mediated by ROS. To address this, we cultured mouse and human PDAC cells with increasing dosages of 3-IAA, 3-IAA and neutrophils or 3-IAA and MPO with or without chemotherapy. 3-IAA directly increased ROS in a dose-dependent manner. The addition of neutrophils or MPO further enhanced the accumulation of ROS in cancer cells, and the addition of MPO reduced the viability of mouse and human PDAC cell lines (Extended Data Fig. 7d–k). Notably, direct addition of the 3-IAA oxidation product MOI also increased the efficacy of FIRINOX in mouse and human PDAC cells (Extended Data Fig. 7l,m).

Furthermore, to test the effects of 3-IAA and FIRINOX treatment on ROS induction in vivo, we treated SPF mice with either 3-IAA and FIRINOX or FIRINOX alone. In line with our in vitro data, treatment with 3-IAA and FIRINOX induced high oxidative stress, measured as ROS in cancer cells by flow cytometry or as nitrotyrosine in whole-tumour histology stains (Extended Data Fig. 7n,o). Further strengthening the link between ROS

production and the oxidation of 3-IAA, we observed much lower levels of ROS after treatment with 3-IAA and FIRINOX in cancer cells from *Mpo*$^{-/-}$ compared to wild-type bone-marrow-reconstituted mice, by flow cytometry (Extended Data Fig. 7p).

We then wondered which oxidative-stress-related pathways are involved in the accumulation of ROS in 3-IAA- and FIRINOX-treated tumours. To this end, we analysed the expression of known ROS-producing or ROS-degrading enzymes in mRNA sequencing data, comparing tumours from mice that were treated with 3-IAA and FIRINOX to tumours from mice treated with FIRINOX alone (Extended Data Fig. 8a). We found that the ROS-degrading enzymes glutathione peroxidase 3 (GPX3) and glutathione peroxidase 7 (GPX7) were downregulated in vivo, and treatment with 3-IAA, FIRINOX and MPO led to a similar downregulation of GPX3 and GPX7 in KPC cells in vitro (Extended Data Fig. 8b). Knockdown of *Gpx3* and *Gpx7* in cancer cells was sufficient to increase the accumulation of ROS after treatment with FIRINOX, and increased the susceptibility of the cancer cells to FIRINOX to a similar extent as 3-IAA and FIRINOX (Extended Data Fig. 8c,d). Of note, knockdown of *Gpx3* was sufficient to establish susceptibility to FIRINOX in vivo (Extended Data Fig. 8e). Finally, we could show that ROS accumulation is essential for the efficacy of 3-IAA and FIRINOX, because treatment with the ROS scavenger *N*-acetylcysteine (NAC) abolished the efficacy of FIRINOX in R-microbiota-colonized (that is, high-3-IAA) mice (Fig. 3a).

Fig. 3: Treatment with 3-IAA and FIRINOX results in reduced autophagic activity.



a, Gnotobiotic mice were colonized with R microbiota and KPC cells were orthotopically injected. Mice were untreated or treated with FIRINOX, NAC (day 9–13) or FIRINOX + NAC ($n = 5$ each). Tumour weight at day 20 of the experiment is shown. **b**, SPF mice bearing orthotopic KPC tumours were substituted +/− 3-IAA, treated with FIRINOX and analysed as indicated. IHC, immunohistochemistry. **c**, Representative images of orthotopic tumours stained with haematoxylin and eosin (H&E), LC3B or Ki67 (left) and respective statistics for positive cells per field ($n = 5$ each) (right). Scale bars, 50 μm. **d**,**e**, The GFP-LC3B-RFP reporter cell line Hy19636_GLRM was injected into SPF mice and mice were treated as indicated ($n = 5$ each). The graphs show the GFP/RFP ratio at day one (**d**) and tumour weight at day three (**e**) after FIRINOX treatment. Representative merged immunofluorescence images or indicated areas with a magnification of 3× are shown; scale bars, 10 μm. **f**, KPC cancer cells were orthotopically injected into R-microbiota-colonized mice and the indicated treatment was applied as shown in the scheme ($n = 4$ or 5). Tumour weight is shown at day 18 of the experiment. **g**, Tumour weight of orthotopic KPC tumours from gnotobiotic mice colonized with NR microbiota is shown nine days after the

indicated treatment (*n* = 4 or 5). CQ, hydroxychloroquine. One experiment (**c**) or one out of two independent experiments (**a**,**d**,**e**–**g**) is shown. Each symbol represents one mouse. Error bars indicate s.e.m. Significant *P* values are indicated and were determined by one-way ANOVA followed by Dunnett's (**a**,**c**,**f**) or Tukey's (**d**,**e**) post-hoc test or Kruskal–Wallis test followed by Dunn's post-hoc test (**g**).

Source data

Next, we wondered what the molecular response to high levels of ROS is. To address this, we performed a targeted protein screen and mRNA sequencing on tumours isolated from SPF mice that were treated with 3-IAA and FIRINOX or FIRINOX alone (Fig. 3b). Autophagy was one of the main pathways that was downregulated in tumours isolated from mice treated with 3-IAA and FIRINOX (Extended Data Fig. 8f,g and Supplementary Tables 1–3). Notably, autophagy is an essential metabolic program for PDAC to proliferate and thrive[29]. Next, we found that the autophagy substrate LC3B was reduced and p62/SQSTM was increased in the tumours isolated from SPF mice that were treated with 3-IAA and FIRINOX (Fig. 3c and Extended Data Fig. 8h). In line with this, tumours from R-microbiota-colonized mice (high 3-IAA) also showed a reduced abundance of LC3B and an increased abundance of p62/SQSTM1 compared to tumours from NR-colonized mice (low 3-IAA) (Extended Data Fig. 8i,j). Changes in LC3B and p62/SQSTM were accompanied by a decline in tumour cell proliferation, measured as the expression of Ki67, in mice treated with 3-IAA and FIRINOX (Fig. 3c). However, we did not observe a change in the number of apoptotic cells, as measured by cleaved caspase-3 (Extended Data Fig. 8k), as late as three days after treatment. To test whether downregulation of autophagy is an essential downstream mechanism that explains the synergy of 3-IAA and FIRINOX, or whether it is merely associated with the observed reduction in cancer cell proliferation, we set up gain- and loss-of-function experiments in vivo. First, we observed that treatment with the disaccharide trehalose[30] normalizes autophagic activity, as measured by a decreasing GFP/RFP ratio in autophagy reporter cells[31] (Fig. 3d and Extended Data Fig. 8l). Second, trehalose was sufficient to completely reverse the treatment efficacy of 3-IAA and FIRINOX, similar to the ROS scavenger NAC (Fig. 3e,f). Third, treatment with the autophagy blocker hydroxychloroquine sensitized

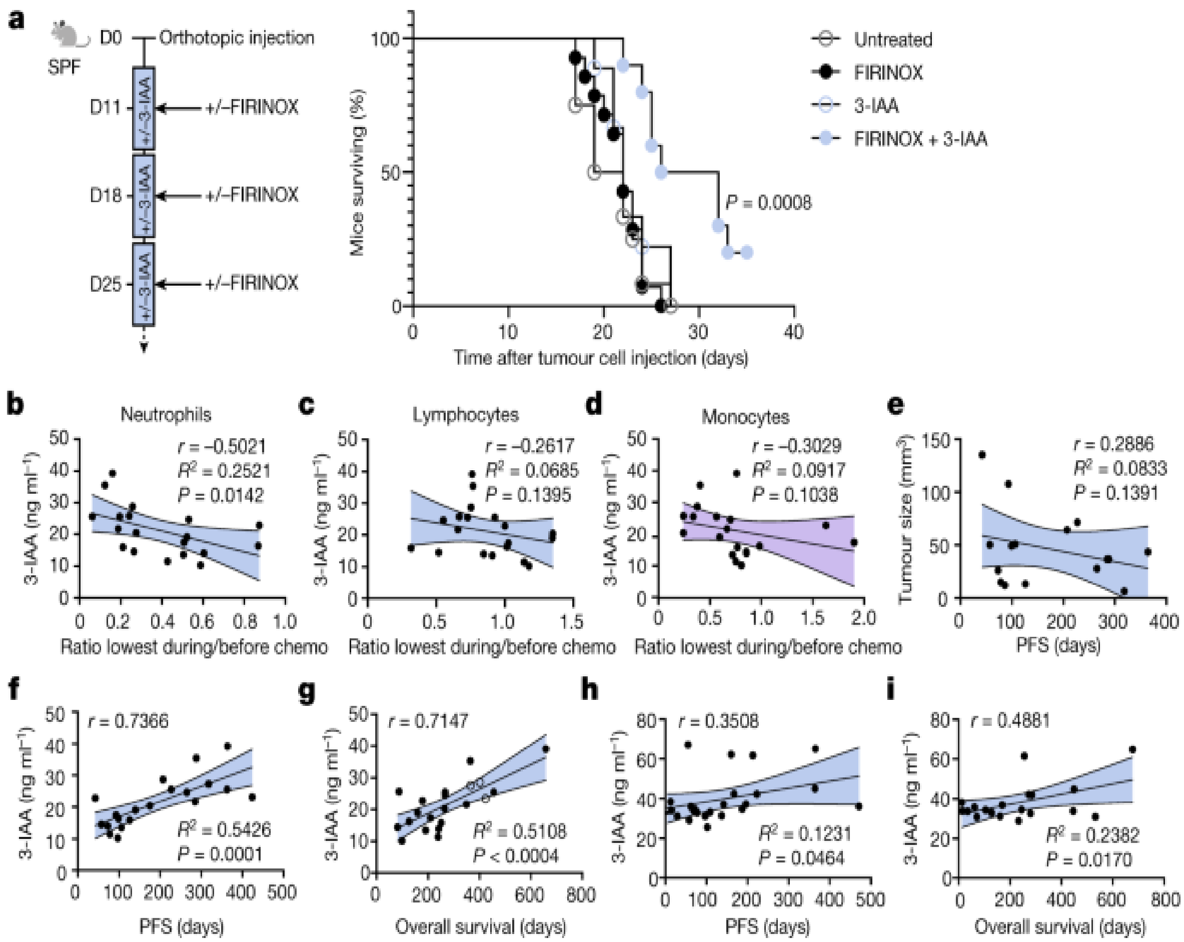
tumours from NR- microbiota-colonized mice (low 3-IAA) to FIRINOX treatment (Fig. 3g). Finally, inhibiting autophagy in cancer cells through a doxycycline-inducible dominant-negative ATG4B protein (mSt-ATG4B)[29] increased the susceptibility to FIRINOX treatment in SPF mice in comparison to control cells (mSt) or mSt-ATG4B cells without doxycycline treatment (Extended Data Fig. 8m).

Overall, these data suggest that when 3-IAA and MPO are present during FIRINOX treatment, the accumulation of ROS increases and the stress adaptation of cancer cells is impaired (that is, autophagy is downregulated), ultimately inducing the reduced proliferation of tumour cells.

3-IAA has therapeutic potential

To further investigate potential therapeutic implications of 3-IAA, we tested the effect of repetitive applications of 3-IAA and FIRINOX on mouse survival, treatment of other cancer entities and the efficacy of 3-IAA with a different chemotherapy combination, namely GnP. Up to three cycles of 3-IAA and FIRINOX together—but neither alone—significantly increased the duration of survival with orthotopic PDAC (Fig. 4a). Notably, 3-IAA was also synergistic with FIRINOX in the treatment of subcutaneous colorectal (MC38) or lung (LLC) tumours (Extended Data Fig. 9a,b) and also synergized with GnP in orthotopic PDAC (Extended Data Fig. 9c). These findings highlight the potential general role of 3-IAA in cancer treatment.

Fig. 4: 3-IAA is clinically relevant in PDAC.



a, SPF mice were orthotopically injected with Hy19636 cells, treated as indicated and their overall survival is depicted in the Kaplan-Meier estimator (untreated *n* = 12; FIRINOX *n* = 14; 3-IAA *n* = 9; 3-IAA + FIRINOX *n* = 10). **b**–**d**, The 3-IAA serum concentration of patients from the Hamburg cohort was measured by chemiluminescence immune assay (CLIA) and correlated with the ratio of blood neutrophil (**b**), lymphocyte (**c**) or monocyte (**d**) counts at the time point of lowest overall leukocyte count (within the first three months of chemotherapy) and counts before start of chemotherapy. **e**, Tumour size as measured in CT scans was correlated with PFS in patients from the Hamburg cohort. **f**,**g**, 3-IAA serum concentration after two to three chemotherapy cycles of patients from the Hamburg cohort was correlated with PFS (**f**) or overall survival (**g**). One patient was excluded from **f**, because the patient's cancer did not progress before the event of death. Patients represented with open circles are still alive and therefore excluded from the correlative analysis in **g**. **h**,**i**, The 3-IAA serum concentration before the start of treatment of patients from the Munich

cohort was correlated with PFS (**h**) or overall survival (**i**). Each symbol represents one patient. Two independent experiments were pooled (**a**). Mean and 95% confidence intervals. *P* values are indicated and were determined by log-rank Mantel–Cox test (**a**) or simple linear regression and Pearson's *r* (**b**–**i**).

Source data

Finally, we wanted to assess the relevance of our findings in humans. We observed higher rates of neutropenia in patients who responded to chemotherapy, consistent with our mouse data and the known association between neutropenia and PDAC prognosis[22] (Extended Data Fig. 9d). Notably, the levels of neutrophil decrease after chemotherapy, but not lymphocyte or monocyte decrease, correlated with 3-IAA serum concentrations, as expected from their different levels of MPO (Fig. 4b–d). Furthermore, we observed a significant correlation of the 3-IAA serum concentration and PFS or overall survival in our observation cohort (Hamburg cohort; Extended Data Table 1), which was absent for the secondary bile acid DCA (Fig. 4f,g and Extended Data Fig. 9e). We validated these findings in a second cohort of patients with mPDAC from the Ludwig-Maximilians-University (LMU) Hospital (Munich cohort; Extended Data Table 1) (Fig. 4h,i). Reflecting the missing value of other clinically assessable markers to predict a chemotherapy response a priori, neither patient-specific (age, gender and weight) nor tumour-specific (tumour size and tumour marker) variables were significantly associated with PFS (Fig. 4e and Extended Data Fig. 9f).

Discussion

In summary, our results identify the microbiota-derived metabolite 3-IAA as a key amplifier of the response to chemotherapy in PDAC. When 3-IAA and MPO are present in high concentrations during FIRINOX treatment, the accumulation of ROS increases and the stress adaptation of cancer cells is impaired, which ultimately results in the reduced proliferation of PDAC cells. Despite the limited number of patients analysed (Hamburg cohort $n =$ 23, of whom 21 had available serum samples; Munich cohort $n = 24$), we observed a robust correlation between 3-IAA serum concentrations

measured during chemotherapy (Hamburg cohort) or even before the start of chemotherapy (Munich cohort) and PFS or overall survival.

These data set the early premise to initiate clinical trials that aim to raise the serum concentration of 3-IAA during chemotherapy treatment through direct treatment or dietary intervention, and ultimately even increase the survival of NR patients with PDAC. Increasing the 3-IAA substrate tryptophan through a specific diet could be easily achieved; however, this approach is affected by the composition of the microbiota, as suggested by the finding that 3-IAA serum concentrations only increased in R-microbiota-colonized mice after dietary intervention. Therefore, when it becomes possible to produce 3-IAA under good manufacturing practice standard conditions, direct treatment with 3-IAA would be ideal, especially considering its capacity to bypass the presence of an undesired microbiota, as shown by our experiments in mice colonized with a NR microbiota.

Before planning future clinical trials, it is important to consider that, in the absence of chemotherapy, indoles or other tryptophan metabolites can impair the development of anti-tumour immune responses through the AhR in PDAC[16]. However, our data show that four to five days of 3-IAA treatment without chemotherapy is not sufficient to affect the growth of the tumour, and only reduces the growth of the tumours via an AhR-independent mechanism when it is combined with chemotherapy. Nevertheless, more observational studies are needed before proceeding with any interventional ones.

Alternatively, faecal microbiota transplantation, which is being tested for the treatment of patients with cancer[8,9], may be sufficient to increase the serum concentrations of 3-IAA and, accordingly, the response to chemotherapy, as shown in our gnotobiotic mice experiments using the stool of ten different human donors. Our study suggests that *B. fragilis* and *B. thetaiotaomicron* are enriched in the microbiota of R patients and are able to produce 3-IAA. However, larger studies are needed to reveal whether these bacteria are indeed responsible for the production of 3-IAA production in R patients, considering the vast number of taxonomically distinct bacterial species that are able to produce 3-IAA (refs. [14,15]). Such studies are of interest to narrow the selection of donors for potential future studies of faecal microbiota transplantation.

In parallel to the effect of tumour size through treatment with 3-IAA and FIRINOX, we also observed a depletion of neutrophils in mouse models of PDAC. In patients with PDAC, we observed a positive correlation between the depletion of neutrophils during chemotherapy and the serum concentration of 3-IAA in the Hamburg cohort. These data suggest that, as in mice, 3-IAA can also contribute to the development of neutropenia in humans. However, given that other factors can lead to neutropenia, such as an altered chemotherapeutic metabolism, blocked excretion of chemotherapeutics or differing chemotherapeutic dosages[32], studies that focus on this particular aspect are needed to firmly conclude whether there is a causative link between 3-IAA and neutropenia in humans. Furthermore, as neutrophil levels are a key predictor of survival in many cancer types, investigations of the described mechanism in other cancers may be warranted[33,34].

Our findings have the potential to be transformative for the treatment of PDAC and other cancer types such as colorectal cancer, especially for patients who are treated with FIRINOX-based regimens. In addition, our findings will trigger the development of studies that address the effects of microbiota-derived metabolites on the ROS–autophagy axis in response to chemotherapeutic treatments.

Methods

Patients and human material

Patients who had been diagnosed with mPDAC and scheduled for treatment with GnP or FOLFIRINOX were recruited to the study ($n = 30$). Informed consent was obtained from all patients as approved by the ethics commission Hamburg (Ethikkommission der Ärztekammer Hamburg, Germany). Patients who received antibiotics during the first three months of treatment, did not deliver pre-treatment stool, did not start chemotherapy or suffered from COVID-19 were excluded from the analyses ($n = 7$). Stool was collected before the start of chemotherapy using a home sampling kit (OMNIgene, Gut OMR-200) for shotgun metagenomic sequencing. One sample was excluded after sequencing owing to overloading of the sample tube that led to unsuccessful fixation and limited representation of the

microbiota. Therefore, the final cohort for microbiota analysis contained 22 patients. The classification of R and NR patients was primarily based on tumour shrinkage of at least 25%, which is calculated as a decrease of the primary tumour and largest metastasis in CT scans comparing the time point before treatment initiation to the time point that revealed the best response during first-line treatment as assessed by the local radiologist (20 to 30% of changes are usually considered significant in CT response evaluation[35]). In cases of disease stabilization or missing CT scans, the criteria of a PFS of over 140 days (median PFS of real-world FOLFIRINOX and GnP cohorts[36]) or a decline of at least 40% in serum tumour markers (prognostic cut-off in palliative PDAC treatment[37]) were considered. Stool for functional experiments was collected in tubes without any additive, transported directly to the laboratory, subsequently diluted in 20% glycerol (Teknova, G1723), aliquoted and frozen at −80 °C. To preserve most of the donor microbiota for functional experiments, the maximum allowed turnaround time from toilet to freezer was 4 h. Blood was drawn at cycle three or four of chemotherapy treatment, mixed 1 to 1 with phosphate-buffered saline (PBS) and plasma was isolated using gradient centrifugation. Serum material of the Munich cohort was taken before the start of chemotherapy and processed according to local standards. Informed consent was obtained from all patients as approved by the ethics commission Munich (project number: 284-10).

Animal models

All mice used in this study were of a C57BL/6 background. All mice were used in accordance with the institutional review board ‘Behörde für Soziales, Familie, Gesundheit und Verbraucherschutz’ (Hamburg, Germany). Mice were kept under SPF or germ-free conditions, with an ambient temperature of 20 ± 2 °C, humidity of 55 ± 10% and a dark–light cycle of 12 h. Age- and sex-matched littermates between 4 and 16 weeks old were used for the most part. *Mpo*$^{-/-}$ bone marrow used to establish bone marrow chimeras was provided by S. Baldus and M. Mollenhauer. *Ahr*$^{-/-}$ bone marrow used to establish bone marrow chimeras was provided by C. Esser.

For colonization of human microbiota, stool was thawed, washed with brain heart infusion broth (BHI) and diluted in BHI. Two hundred microlitres of

the suspension was gavaged once orally to gnotobiotic mice housed in isocages. Two to four weeks later, tumour experiments were initiated.

Sample sizes were calculated on the basis of small pilot experiments and mice were randomized before the beginning of treatment. The person treating the mice was not blinded because of the complex treatment schedules. For models of orthotopic PDAC, 5×10^4–10×10^4 KPC, 2×10^5 Hy19636, 1×10^5 mSt-ATG4B/mSt or 1×10^5 knockdown KPC (AhR, GPX, scramble control) cells were injected orthotopically in a 1:1 mixture of PBS and Matrigel (Corning, 356231). Tumour weight was assessed at day 20 after tumour cell injection unless otherwise indicated. For subcutaneous tumour growth, 2.5×10^5 LLC-GFP or 1×10^6 MC38 cells were subcutaneously injected into the flank. Subcutaneous tumour growth was measured every other day using a caliper. Tumour weight was assessed at day 17 after tumour cell injection. The maximum allowed diameter of subcutaneous tumours for active experimental mice was 1.5 cm and limits were not exceeded.

For bone marrow chimeras, recipient mice were irradiated with 9.6 Gy (BioBEAM 2000) 24 h before bone marrow transfer. One day later, 1×10^6–4×10^6 bone marrow cells isolated from wild-type, *Ahr*$^{-/-}$ or *Mpo*$^{-/-}$ mice were transferred by intravenous injection. Three to four weeks after transfer, cancer cells were orthotopically injected and treated as described below. Successful engraftment was validated by quantitative PCR (qPCR) on tumour-infiltrating immune cells.

Treatment with chemotherapy was initiated intraperitoneally (i.p.) at the indicated time point (usually day 11 after cancer cell injection). Oxaliplatin (Accord) 5 mg kg^{-1}, irinotecan (Accord) 20 mg kg^{-1} and 5-FU (Medac) 50 mg kg^{-1} were used for FIRINOX treatment as described in other studies[38,39]. Folinic acid was not used and gemcitabine (Hexal) 120 mg kg^{-1} and nab-paclitaxel (Celgene) 30 mg kg^{-1} were used for GnP treatment. Depletion of T cells was achieved by treatment with 200 µg anti-CD8 antibody clone 53-6.7 (BioXcell, BE0004) alone or in combination with 200 µg anti-CD4 antibody clone GK1.5 (BioXcell, BE0003) every third day i.p., beginning one day before chemotherapy treatment. Control mice were treated with the respective isotype control clone 2A3 (BioXcell, BE0089) at similar

concentrations and intervals. Autophagy induction was achieved using two i.p. injections of trehalose (Sigma, T9449) at a concentration of 3 g kg^{-1} one day before and on the day of chemotherapy treatment. Similarly, 60 mg kg^{-1} chloroquine (Sigma, C6628) dissolved in PBS was injected i.p. one day before and on the day of chemotherapy treatment.

Treatment with 3-IAA (500 mg kg^{-1}) was applied by oral gavage using indole-acetic-acid sodium salt (Sigma, I5148) dissolved in PBS every day for five consecutive days (two days before chemotherapy until two days after chemotherapy), unless otherwise indicated. Indole-3-propionic acid (3-IPA; Sigma, 57400), GCA (Sigma, G2878), hippuric acid (Sigma, 112003) and DCA (Sigma, 30960) were dissolved in 1 M NaOH in PBS and pH-adjusted to 7.4 using 1 M HCl. The solution was gavaged for five consecutive days orally at a concentration of 500 mg kg^{-1} (3-IPA, hippuric acid) or 250 mg kg^{-1} each (GCA and DCA). NAC (Sigma, a7250) was applied for five days in the drinking water ad libitum at a concentration of 1 g l^{-1}.

Dietary tryptophan modulation was initiated three days before chemotherapy until one day after treatment. In one experiment, tryptophan modulation was applied for a total of 14 days. Standard diet (2.3 g kg^{-1} tryptophan; Altromin, 1320) was changed to either synthetic crystalline AA tryptophan-free (0 g kg^{-1}; SSNIFF, S9336-E701) diet or crystalline AA tryptophan-high (12 g kg^{-1}; SSNIFF, S5714-E711) diet. Subsequently, the diet was changed back to the standard diet.

Dietary doxycycline (Sigma, D9891) was administered at a dosage of 625 mg kg^{-1} through the diet starting at day five or eight after cancer cell injection for a total of up to seven consecutive days. Subsequently, the diet was changed back to the standard diet.

For measurement of serum metabolites, blood was drawn at the end of the experiment or at the indicated time point. Blood was allowed to clot for 30 min and was centrifugated (1,000*g*) for 10 min thereafter. Serum was diluted and used as described in the specific section.

DNA extraction and shotgun metagenomics

For DNA extraction, samples were isolated with the ZymoBIOMICS 96 MagBead DNA kit (Zymo Research, D4302) and purified with Zymo DNA Clean and Concentrator-5 (Zymo Research, D4004). Libraries were prepared with the NEBnext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs, E7645) with 150 ng of total DNA, size selection of 400–500 bp and 4× PCR cycles.

For shotgun metagenomics, Illumina library preparation was performed using the NEBNext Ultra II FS DNA Library Prep Kit (New England Biolabs, E7805). The library preparation was performed according to the manufacturer's instructions. The size selection was performed using AMPure XP beads (Beckman Coulter, A63882) and adaptor enrichment was performed using seven cycles of PCR using the NEBNext Multiplex oligos (New England Biolabs, E7335) from Illumina and then subjected to Illumina NovaSeq (2 × 150 bp) sequencing.

For bioinformatic analysis, raw reads were trimmed for low quality and filtered against the phix174 and human hg19 genome with bbduk (ref. [40]). For taxonomic species profiling, all libraries were mapped against the Unified Human Gastrointestinal Genome collection ($n = 4,644$) using BBMap (refs. [41,42]). Mapping rates were normalized into transcripts per million (TPM) and genomes with less than 10% genome coverage (genome breadth) were considered not prevalent in the sample. Data were summarized as metagenomics operational taxonomic units (OTUs) into biom format and analysed with phyloseq and LEfSe (refs. [43,44]). Species-level functional profiling was performed with HUMAnN3 and also using the 9.9 million gene integrated reference catalogue of the human microbiome[41].

Bacterial strains and isolation

Bacteroides thetaiotaomicron (DSM 2079), *B. fragilis* (DSM 2151) and *Prevotella copri* (DSM 18205) were obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ). For bacteria isolation, faecal samples frozen in glycerol were thawed and streaked out anaerobically in serial dilutions on BHI blood agar plates (5% defibrinated sheep's blood & vitamin K_3) supplemented with vancomycin. After growing the agar plates inside an incubator at 37 °C for 2 days, single colonies were picked into

BHI-S medium in a 96-well plate and were further incubated for 24 h. The resulting bacterial cultures were screened by PCR using specific primers for *B. thetaiotaomicron* (*B. theta* F 5′-GAGGGTGTCGTATTTCCGAAGG-3′ R 5′-GTTCCCTGATCCAGTGTGTTGG-3′) or *B. fragilis* (*B. frag* F 5′-AATGATTCCGCATGGTTTCA-3′ R 5′-ATTTTGGGATTAGCATACGG-3′). After passaging PCR-positive wells on agar plates to obtain pure cultures and additional confirmation of identity by Sanger sequencing, resulting strains were maintained in BHI-S until further use. All bacterial strains were cultured and maintained anaerobically in BHI broth supplemented with 10% fetal bovine serum (FBS) and vitamin K_3 (BHI-S).

Extraction of bacterial supernatant

BHI-S medium supplemented with 1% tryptophan was inoculated from a fully grown overnight bacterial culture (1:50 ratio) and was incubated anaerobically until the early exponential phase. Cultures were taken out of the anaerobic chamber and were centrifuged at 4,700 rpm (4,816*g*) for 5 min at room temperature. The supernatant was removed and immediately frozen at −20 °C.

16S rRNA sequencing

DNA from tumours was extracted using the DNeasy Blood & Tissue Kit (Qiagen, 69504). Approximately 10 mg tissue was digested with Proteinase K in ATL buffer at 56 °C for 1 h. Afterwards, samples were processed according to the manufacturer’s protocol. Blank extraction controls were included during extraction of samples.

Variable regions V1 and V2 of the 16S rRNA gene were amplified using the primer pair 27F-338R in a dual-barcoding approach according to a previous report[45]. For tumours, 3.5 µl DNA was used for amplification and PCR products were verified using agarose gel electrophoresis. Final PCR products were normalized using the SequalPrep Normalization Plate Kit (Thermo Fisher Scientific, A1051001), pooled equimolarly and sequenced on the Illumina MiSeq v3 2×300bp (Illumina). Demultiplexing after sequencing was based on 0 mismatches in the barcode sequences. Data processing was performed using the DADA2[46] workflow for big datasets

(https://benjjneb.github.io/dada2/bigdata.html; the workflow adjusted for the V1–V2 region can be found here: https://github.com/mruehlemann/ikmb_amplicon_processing/blob/master/dada2_16S_workflow.R), resulting in abundance tables of amplicon sequence variants (ASVs). ASVs underwent taxonomic annotation using the Bayesian classifier provided in DADA2 and using the Ribosomal Database Project (RDP) version 16 release. Sequences that were not assignable to genus level were binned into the finest possible taxonomic classification.

Metabolomic screen

We used ultra-high-pressure liquid chromatography–tandem mass spectrometry (UPLC–MS/MS) to acquire data in both positive and negative ionization modes that allowed for identification and quantification of 630 metabolites. Plasma samples were processed using the MxP Quant 500 Kit (Biocrates) according to the manufacturer's instructions. In brief, 10 µl of plasma sample, calibration standard and control sample were transferred onto a filter containing internal standards for internal standard calibration. Filters were dried under a stream of nitrogen using a pressure manifold (Waters). Samples were incubated with derivatization reagent phenyl isocyanate for 60 min. After drying under nitrogen, analytes were extracted with 5 mmol l^{-1} ammonium acetate in methanol and the eluate was further diluted for the UPLC–MS/MS analysis. The targeted analysis covered 630 metabolites (https://biocrates.com/mxp-quant-500-kit/) detected by MS/MS after UPLC separation and flow injection analysis (FIA). Each measurement required two UPLC runs and three FIA runs to cover all metabolites. All analyses were performed on an ACQUITY UPLC I-Class system (Waters) coupled to a Xevo TQ-S mass spectrometer (Waters). Reversed-phase chromatographic separation was accomplished using a C18 LC-column (Biocrates) with 0.2% formic acid in water with 0.2% formic acid in acetonitrile as the eluent system. The FIA solvent was methanol, with a modifier provided by the kit manufacturer. Data analysis of the UPLC–MS/MS results was based on a seven-point curve or one-point calibration and internal standard normalization. Values below the lower threshold were set to zero. Concentration data were analysed using MetaboAnalyst v.5. Concentrations were log-transformed before analysis and raw *P* values and $\log_2$-transformed fold change values are shown in the graphics of Fig. 1.

3-IAA and DCA CLIA

To quantify 3-IAA or DCA serum concentrations, mouse or human serum was diluted 1 to 10 with PBS and the chemiluminescence immune assay (CLIA) (Abbexa, 3-IAA abx190011; DCA 258844) was performed according to the manufacturer's protocol. For detection of 3-IAA in cultures, supernatant was processed as described above and directly used for the assay. Chemiluminescence was detected using FLUOstar Omega (BMG Labtech) with 1-s sampling per well, and a gain of 3,400 to 4,000 was individually adjusted for each assay. The concentration was determined using supplied standards and interpolated using Prism 8.4.0.

3-IAA and MOI measurement in tumour tissue

Frozen tumour tissue was mixed with ice-cold methanol 1:1–3 and homogenized using 0.4-mm beads at 5,500 rpm for 2 × 30 s using a homogenizer (Precyllys 24 touch). Extracts were centrifuged (10,000*g*) at 4 °C and the supernatant was used for LC–MS/MS analysis. Reversed-phase chromatography was conducted using a biphenyl stationary phase (Raptor Biphenyl (Restek), dimensions: 50 mm × 2.1 mm ID; particle size: 2.7 µm) with eluent A: water + 0.1% formic acid + 5 mM ammonium acetate and eluent B: methanol + 0.1% formic acid + 5 mM ammonium acetate. The flow rate was set to 0.5 ml per min. Elution starts with 95% eluent A, which linearly decreases to 75% over 0.5 min. This composition was held for 4 min before returning to initial conditions. The injected volume was 2 µl and the column temperature was set to 55 °C. 3-IAA and MOI were detected in MRM (multiple reaction monitoring) mode. The following transitions were monitored after positive electrospray ionization: 3-IAA: *m*/*z* 176.2 > 103.0; *m*/*z* 176.2 > 130.2; MOI: *m*/*z* 148.1 > 120.2; *m*/*z* 148.1 > 130.2; *m*/*z* 148.1 > 133.1. Quantification was done according to a standard curve.

Isolation of immune cells and flow cytometry

Tumours were taken and cut into similar-sized pieces. Tumours were rinsed with cold PBS and digested in RPMI (Sigma, 61870044) supplemented with 10% FBS (Gibco, 10500064), 2.5 mg ml^{-1} collagenase D (Roche, 11088866001) and 0.2 mg ml^{-1} DNase I (Roche, 11284932001) for 35 min

at 37 °C with continuous shaking. Afterwards, the suspension was strained through a 40-μm cell strainer and quenched with cold PBS. Subsequently, immune or tumour cells were stained with Fc block and live/dead staining (Thermo Fisher Scientific L34957 and L10119) for 30 min in the dark. Afterwards cells were washed, stained with the indicated flow cytometry antibodies and incubated for 30 min in the dark again. Flow cytometry was performed on a Fortessa flow cytometer (BD). To assess the cytokine profile of immune cells, restimulation of T cells with 50 ng ml^{-1} PMA, 500 ng ml^{-1} ionomycin and 1 μg ml^{-1} brefeldin A was performed for 3 h at 37 °C. After surface staining, cells were fixed and permeabilized using the eBioscience Foxp3 intracellular staining kit (00-5523-00). The following intracellular antibodies were used: IFNγ (1:200 dilution) and TNF (1:400 dilution). The following surface antibodies were used to classify lymphocytes (CD3 (1:300 dilution), CD4 (1:500 dilution), CD8 (1:400 dilution), CD19 (1:100 dilution), CD44 (1:500 dilution), PD-1 (1:400 dilution) and NK1.1 (1:400 dilution)) or myeloid cells (CD11b (1:800 dilution), CD11c (1:300 dilution), CD45 (1:800 dilution), Ly6G (1:800 dilution), Ly6C (1:600 dilution), Ly6B (1:200 dilution), CD115 (1:300 dilution), F4/80 (1:600 dilution) and MHCll (1:600 dilution)). Epithelial cancer cells were stained with EPCAM (1:200 dilution) antibody. For Ly6B staining, the goat anti-rat IgG antibody (1:400 dilution) was used. Cells were counted by flow cytometry using 2,500 to 5,000 beads (Spherotec, ACBP 100-10). Viability of neutrophils in vitro was assessed using PI/Annexin V staining according to the manufacturer's protocol. ROS expression was determined in immune cells in vivo after tissue digestion as described above. Afterwards, the ROS dye CellROX (Thermo Fisher Scientific, C10422) was stained in parallel with flow cytometry antibodies at a dilution of 1:1,000. Software analysis and histogram generation was carried out using FlowJo v.10.

Cell culture

KPC cells were obtained from Ximbio under catalogue number 153474; Hy19636_GLRM reporter cells and mSt-ATG4B/mSt cells were provided by A. Kimmelman[31]; MC38 and LLC-GFP (ATCC) cells were provided by A. Giannou; and MIA PaCa-2, BxPC-3 and T3M-4 cells (all ATCC) were provided by C. Güngör. All cells tested negative for mycoplasma contamination. Cells were maintained under standard conditions at 37 °C

and 5% CO_2 and regularly visually inspected. Cells were grown in DMEM GlutaMAX (Thermo Fisher Scientific, 10566016) supplemented with penicillin and streptomycin (Gibco, 15140122) and 10% FBS (Gibco, 10500064). For in vitro experiments, cancer cells were plated at 5,000–10,000 cells per well in a 96-well plate. Cells were allowed to seed overnight, except in experiments with neutrophil co-culture, treatment of GPX-knockdown cells or co-treatment with MPO. Subsequently, treatment with indicated compounds and treatment duration was initiated. 3-IAA or 3-IPA (3-IPA, Sigma, 57400; 3-IAA, Sigma, I3750) were dissolved in 1 M NaOH in PBS and PH-adjusted to 7.4 using 1 M HCl or DMSO when indicated. NAC was diluted in PBS and used at a concentration of 1 mM. H_2O_2 (Sigma, H1009) was used at a concentration of 400 μM.

Tumour cell viability and proliferation was assessed using a MTT or MTS assay (Abcam, ab201191 and ab197010) according to the manufacturer's protocol. Absorbance was assessed using the FLUOstar Omega (BMG Labtech). In other experiments, viability was assessed by flow cytometry using SYTOX (Thermo Fisher Scientific, S34857), PI (Biolegend, 421301) and 2,500–5,000 counting beads (Spherotec, ACBP 100-10) as a reference. In parallel, intracellular ROS was measured using CellROX (Thermo Fisher Scientific, C10422) according to the manufacturer's protocol and assessed by flow cytometry.

Neutrophils or T cells were isolated from bone marrow or spleens of healthy, untreated mice and sorted using fluorescence-activated cell sorting (FACS) as $LY6G^+CD11b^+$ and $TCR\beta^+CD4^+$ or $CD8^+$, respectively. Cells were seeded at a concentration of 20,000–50,000 per well in a 96-well plate. Viability of neutrophils or T cells was assessed by flow cytometry using SYTOX (Thermo Fisher Scientific, S34857) or PI/Annexin V staining (Biolegend, 640914) at the indicated time of the experiment. In some experiments, 200 mU ml^{-1} MPO or 400 mU ml^{-1} MPO (Merck, 475911) or indicated concentrations of MOI (Sigma, 493397) were added. Degranulation of neutrophils was assessed by the release of MPO. A total of 1×10^6 neutrophils were sorted using FACS as described above. Neutrophils were incubated in HBSS (Gibco, 14065-56) and the indicated treatment, or *N*-formylmethionyl-leucyl-phenylalanine (fMLP; Merck, F3506), as a positive control, was added. After 30 min of incubation, MPO activity was

measured using an MPO activity assay kit (Abcam, ab105136) according to the manufacturer's protocol. NET formation of neutrophils was determined using SYTOX stain after 3 h of incubation with the indicated compounds or 100 nM phorbol-12-myristate-13-acetate (PMA; Merck, 524400) as a positive control. NETs were measured as SYTOX-positive cells using flow cytometry.

Pre-neutrophils were sorted by FACS (lineage-negative $(CD3,NK1.1,CD19,B220)^{-}$, $CD115^{-}$, $Ly6B^{+}$, $Ly6G^{int\text{-}low}$) as previously described[26] and cultured at a density of 20,000–50,000 per well in a 96-well plate. Indicated treatments were applied and viability was assessed using PI staining in flow cytometry.

MPO activity

Intratumoural or bone-marrow-derived neutrophils ($LY6G^{+}CD11b^{+}$) or pre-neutrophils (lineage-negative $(CD3,NK1.1,CD19,B220)^{-}$, $CD115^{-}$, $Ly6B^{+}$, $Ly6G^{int\text{-}low}$) were sorted by FACS. A total of 50,000 cells were processed for MPO activity measurement using an MPO activity assay kit according to the manufacturer's protocol (Abcam, ab219925). Fluorescence was analysed using FLUOstar Omega (BMG Labtech). MPO activity was calculated based on a standard curve as suggested by the manufacturer's protocol.

Protein screen

Tumour tissue from three individual mice per group was pooled and proteins were extracted using scioExtract buffer (Sciomics). The samples were labelled for 2 h with scioDye 2 (Sciomics) at an adjusted protein concentration. The reference sample was labelled with scioDye 1 (Sciomics). After 2 h, the reaction was stopped and the buffer was exchanged with PBS. All labelled protein samples were stored at −20 °C until use. The samples were analysed in a dual-colour approach using a reference-based design on scioDiscover antibody microarrays (Sciomics). The arrays were blocked with scioBlock (Sciomics) on a Hybstation 4800 (Tecan) and afterwards the samples were incubated competitively with the reference sample using a dual-colour approach. After incubation for 3 h, the slides were thoroughly washed with 1× PBSTT, rinsed with 0.1× PBS as

well as with water and subsequently dried with nitrogen. Slide scanning was conducted using a Powerscanner (Tecan) with constant instrument laser power and PMT settings. Spot segmentation was performed with GenePix Pro 6.0 (Molecular Devices). Acquired raw data were analysed using the (LIMMA) package of R-Bioconductor after uploading the median signal intensities. For normalization, a specialized invariant Lowess method was applied. Downregulated proteins (M-value of < -0.35) or upregulated proteins (>0.35) in the 3-IAA and FIRINOX sample were uploaded to the STING database (https://string-db.org/cgi/input.pl). The standard pipeline was used to perform KEGG enrichment analysis. Only pathways with a false discovery rate (FDR) < 0.05 were considered statistically significant.

RNA extraction and mRNA sequencing

Fresh frozen tumour tissue was lysed in TRIzol reagent (Thermo Fisher Scientific, 15596018) and RNA was extracted using the chloroform-isopropanol method. mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. After fragmentation, the first-strand cDNA was synthesized using random hexamer primers followed by second-strand cDNA synthesis. The library was ready after end repair, A-tailing, adapter ligation, size selection, amplification and purification. The library was checked with Qubit and qPCR for quantification, and Bioanalyzer for size distribution detection. Quantified libraries were sequenced on the Illumina platform with at least 60 million reads per sample.

Sequence reads were processed with fastp (v0.20.1) to remove sequences of sequencing adapters and low-quality (Phred quality score below 15) sequences from the 3′ end of the sequence reads. Afterwards, reads were aligned to the mouse reference assembly (GRCm39.104) using STAR (v.2.7.9a)[47]. Differential expression was assessed with DESeq2 (ref. [48]). A gene was considered significantly differentially expressed if the corresponding absolute $\log_2$-transformed fold change (log2FC) was no less than 0.6 and, in addition, the FDR did not exceed a value of 0.1. The gene set enrichment analysis (GSEA) of the Reactome pathway Autophagy (R-HSA-9612973) was performed using fgsea (v.4.1)[49].

qPCR

Total RNA was extracted from cell lines using Trizol Reagent (Invitrogen,15596018) and the total RNA extraction kit (Qiagen, 74004/74104) according to the manufacturer's protocol. The High-Capacity cDNA Synthesis Kit (Thermo Fisher Scientific, 4368813) was used for cDNA synthesis. Primers and probes were purchased from Applied Biosystems. Mouse primers and probes are: *Gpx3* (Mm00492427_m1), *Gpx7* (Mm00481133), *Ahr* (Mm00478932_m1). qPCR was performed using the TaqMan Master Mix (Thermo Fisher Scientific, 4369016) on the StepOne Plus system (Applied Biosystems). Forty to 44 cycles were applied for every assay and technical doublets or triplicates were used. If at least two out of three values of the technical replicates were undetectable, the expression was considered non-detectable. Relative expression was normalized to *Gapdh* (Mm99999915_g1).

Lentiviral transfer of shRNAs

Lentiviral vectors expressing short hairpin RNAs (shRNAs) under control of the human U6 promoter (MISSION pLKO.1-puro) directed against mouse *Gpx3* (TRCN000076539), mouse *Gpx7* (TRCN0000076563) and mouse *Ahr* (TRCN0000218025) as well as a nontargeted control shRNA (SHC002, scrambled) were obtained from Sigma-Aldrich. The production of lentiviral particles has been described in detail elsewhere[50], and protocols are available online (http://www.LentiGO-Vectors.de). For the transduction of KPC cells with the HIV-1-derived lentiviral vectors, 2.5×10^4 cells were plated in 0.5 ml medium with 8 μg ml^{-1} polybrene per well of a 24-well plate. After plating, the addition of 10 μl VSV-G pseudotyped, non-concentrated lentiviral particles led to the stable integration of shRNAs and the puromycin resistance gene into the cell's genome. To increase the transduction rate by spin-inoculation, the plate was centrifuged at 1,000*g* and 25 °C for 1 h. The selection of successfully transduced cells with 1 μg ml^{-1} puromycin in the culture medium was started four days after transduction.

Immunohistochemistry and quantification

Tissues were fixed in 4% formalin in PBS and embedded in paraffin using an ASP300S dehydration machine (Leica) and an EG1160 tissue embedding

system (Leica). Paraffin sections (2 µm) were cut and stained with H&E or processed for immunohistochemistry as follows: after dewaxing and inactivation of endogenous peroxidases (3% hydrogen peroxide in PBS), antibody-specific heat-mediated antigen retrieval was performed using the Ventana Benchmark XT machine (Ventana). Sections were blocked (10% FCS in PBS) and then incubated with anti-LC3B antibody (1:400 dilution, Thermo Fisher Scientific, PA1-46286); anti-nitrotyrosine antibody (1:100 dilution, Thermo Fisher Scientific, A-21285); anti-CC3 antibody (1:100 dilution, Cell Signaling, 9661); p62/SQSTM1 (1:500 dilution, Thermo Fisher Scientific, PA5-20839); and anti-Ki67 antibody (1:100 dilution, Abcam, 15580). For detection of specific binding, the Ultra View Universal DAB Detection Kit (Ventana, Roche) was used, which contains secondary antibodies, DAB stain and haematoxylin counterstaining reagent. Slides were scanned using NanoZoomer 2.0-HT (Hamamatsu Photonics) and representative images were taken using Fiji.

Quantification for LC3B, Ki67, nitrotyrosine and CC3 was done in a blinded manner. Positive cells were determined using ImageJ v.2.1.0/1.53c and the threshold was adjusted according to the staining intensity of the respective antibody and maintained for all tumours analysed with the same staining. The number of positive cells was counted per 250–500 × 500-µm field (10× to 40× magnification) in three to five fields per sample.

Widefield microscopy

To visualize the GFP-LC3-RFP reporter signal expression of Hy19636_GLRM PDAC cells, tissues were fixed in 2% PFA solution at 4 °C overnight, incubated in PBS containing 30% sucrose and embedded in Tissue-Tek OCT compound (Sakura Finetek) on dry ice. For further analysis, 7-µm sections were used. Widefield imaging was performed using the THUNDER Imager 3D Live Cell and 3D Cell Culture (Leica Microsystems) equipped with a 40× 1.10 NA water immersion objective. LED power and exposure time and other system-specific settings were first optimized using positive control tissue and not changed between image acquisitions of the different groups to provide optimal comparability. For each tissue section, an average of five regions of interest were randomly selected and imaged for subsequent quantitative analysis. ImageJ imaging

software was used for file navigation, adjustment of colour balance and image analysis. GFP and RFP quantification was determined using ImageJ v.2.1.0/1.53c. The mean fluorescence intensity for each respective signal was determined per slide in at least five areas per tumour.

Graphical abstracts

Figure 1a, Extended Data Fig. 2a,d and small icons in Figs. 1–4 were created with BioRender.com.

Statistical analyses

All statistical analyses, unless otherwise indicated, were performed using GraphPad Prism 9.3.1. Normality and log-normality were tested using Shapiro–Wilk or Kolmogorov–Smirnov tests. If normality was not given, non-parametric testing was performed. Tests were performed two-sided if not otherwise indicated and resulting significant ($P < 0.05$) P values are shown.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

RNA-seq data have been submitted to the European Nucleotide Archive (ENA). They are publicly available under accession number PRJEB58222. Shotgun metagenomic sequencing data were filtered for human reads and are available under accession number PRJEB58222. Source data are provided with this paper.

References

1. Nevala-Plagemann, C., Hidalgo, M. & Garrido-Laguna, I. From state-of-the-art treatments to novel therapies for advanced-stage pancreatic

cancer. *Nat. Rev. Clin. Oncol.* **17**, 108–123 (2020).

2. Rahib, L., Wehner, M. R., Matrisian, L. M. & Nead, K. T. Estimated projection of US cancer incidence and death to 2040. *JAMA Netw. Open* **4**, e214708 (2021).

3. Von Hoff, D. D. et al. Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. *N. Engl. J. Med.* **369**, 1691–1703 (2013).

4. Conroy, T. et al. FOLFIRINOX or gemcitabine as adjuvant therapy for pancreatic cancer. *N. Engl. J. Med.* **379**, 2395–2406 (2018).

5. Raghavan, S. et al. Microenvironment drives cell state, plasticity, and drug response in pancreatic cancer. *Cell* **184**, 6119–6137 (2021).

6. Aung, K. L. et al. Genomics-driven precision medicine for advanced pancreatic cancer: early results from the COMPASS trial. *Clin. Cancer Res.* **24**, 1344–1354 (2018).

7. McQuade, J. L., Daniel, C. R., Helmink, B. A. & Wargo, J. A. Modulating the microbiome to improve therapeutic response in cancer. *Lancet Oncol.* **20**, e77–e91 (2019).

8. Davar, D. et al. Fecal microbiota transplant overcomes resistance to anti-PD-1 therapy in melanoma patients. *Science* **371**, 595–602 (2021).

9. Baruch, E. N. et al. Fecal microbiota transplant promotes response in immunotherapy-refractory melanoma patients. *Science* **371**, 602–609 (2021).

10. Spencer, C. N. et al. Dietary fiber and probiotics influence the gut microbiome and melanoma immunotherapy response. *Science* **374**, 1632–1640 (2021).

11. Riquelme, E. et al. Tumor microbiome diversity and composition influence pancreatic cancer outcomes. *Cell* **178**, 795–806 (2019).

12. Balachandran, V. P. et al. Identification of unique neoantigen qualities in long-term survivors of pancreatic cancer. *Nature* **551**, 512–516

(2017).

13. Thomas, R. M. & Jobin, C. Microbiota in pancreatic health and disease: the next frontier in microbiome research. *Nat. Rev. Gastroenterol. Hepatol.* **17**, 53–64 (2020).

14. Russell, W. R. et al. Major phenylpropanoid-derived metabolites in the human gut can arise from microbial fermentation of protein. *Mol. Nutr. Food Res.* **57**, 523–535 (2013).

15. Kaur, H., Bose, C. & Mande, S. S. Tryptophan metabolism by gut microbiome and gut–brain-Axis: an in silico analysis. *Front. Neurosci.* **13**, 1365 (2019).

16. Hezaveh, K. et al. Tryptophan-derived microbial metabolites activate the aryl hydrocarbon receptor in tumor-associated macrophages to suppress anti-tumor immunity. *Immunity* **55**, 324–340 (2022).

17. Opitz, C. A. et al. The therapeutic potential of targeting tryptophan catabolism in cancer. *Br. J. Cancer* **122**, 30–44 (2020).

18. Gál, E. et al. Bile accelerates carcinogenic processes in pancreatic ductal adenocarcinoma cells through the overexpression of MUC4. *Sci. Rep.* **10**, 22088 (2020).

19. Yoshimoto, S. et al. Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome. *Nature* **499**, 97–101 (2013).

20. Reyes, C. M. et al. Neoadjuvant therapy remodels the pancreatic cancer microenvironment via depletion of protumorigenic immune cells. *Clin. Cancer Res.* **26**, 220–231 (2020).

21. Pushalkar, S. et al. The pancreatic cancer microbiome promotes oncogenesis by induction of innate and adaptive immune suppression. *Cancer Discov.* **8**, 403–416 (2018).

22. Roviello, G. et al. Association between neutropenia and survival to nab-paclitaxel and gemcitabine in patients with metastatic pancreatic cancer. *Sci. Rep.* **10**, 19281 (2020).

23. De Melo, M. P., Curi, T. C. P., Miyasaka, C. K., Palanch, A. C. & Curi, R. Effect of indole acetic acid on oxygen metabolism in cultured rat neutrophil. *Gen. Pharmacol.* **31**, 573–578 (1998).

24. Folkes, L. K., Dennis, M. F., Stratford, M. R. L., Candeias, L. P. & Wardman, P. Peroxidase-catalyzed effects of indole-3-acetic acid and analogues on lipid membranes, DNA, and mammalian cells in vitro. *Biochem. Pharmacol.* **57**, 375–382 (1999).

25. Folkes, L. K., Rossiter, S. & Wardman, P. Reactivity toward thiols and cytotoxicity of 3-methylene-2-oxindoles, cytotoxins from indole-3-acetic acids, on activation by peroxidases. *Chem. Res. Toxicol.* **15**, 877–882 (2002).

26. Kim, M. H. et al. A late-lineage murine neutrophil precursor population exhibits dynamic changes during demand-adapted granulopoiesis. *Sci. Rep.* **7**, 39084 (2017).

27. Stockinger, B., Di Meglio, P., Gialitakis, M. & Duarte, J. H. The aryl hydrocarbon receptor: multitasking in the immune system. *Annu. Rev. Immunol.* **32**, 403–432 (2014).

28. Perillo, B. et al. ROS in cancer therapy: the bright side of the moon. *Exp. Mol. Med.* **52**, 192–203 (2020).

29. Yang, A. et al. Autophagy sustains pancreatic cancer growth through both cell-autonomous and nonautonomous mechanisms. *Cancer Discov.* **8**, 276–287 (2018).

30. DeBosch, B. J. et al. Trehalose inhibits solute carrier 2A (SLC2A) proteins to induce autophagy and prevent hepatic steatosis. *Sci. Signal.* **9**, ra21 (2016).

31. Yamamoto, K. et al. Autophagy promotes immune evasion of pancreatic cancer by degrading MHC-I. *Nature* **581**, 100–105 (2020).

32. Svedberg, A. et al. Genetic association of gemcitabine/carboplatin-induced leukopenia and neutropenia in non-small cell lung cancer

patients using whole-exome sequencing. *Lung Cancer* **147**, 106–114 (2020).

33. Rossini, D. et al. Clinical impact of neutropenia and febrile neutropenia in metastatic colorectal cancer patients treated with FOLFOXIRI/bevacizumab: a pooled analysis of TRIBE and TRIBE2 studies by GONO. *ESMO Open* **6**, 100293 (2021).

34. Gargiulo, P. et al. Chemotherapy-induced neutropenia and treatment efficacy in advanced non-small-cell lung cancer: a pooled analysis of 6 randomized trials. *BMC Cancer* **21**, 549 (2021).

35. Hayes, S. A. et al. Comparison of CT volumetric measurement with RECIST response in patients with lung cancer. *Eur. J. Radiol.* **85**, 524–533 (2016).

36. Riedl, J. M. et al. Gemcitabine/nab-paclitaxel versus FOLFIRINOX for palliative first-line treatment of advanced pancreatic cancer: a propensity score analysis. *Eur. J. Cancer* **151**, 3–13 (2021).

37. Koom, W. S., Seong, J., Kim, Y. B., Pyun, H. O. & Song, S. Y. CA 19-9 as a predictor for response and survival in advanced pancreatic cancer patients treated with chemoradiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* **73**, 1148–1154 (2009).

38. Grierson, P. M. et al. The MK2/Hsp27 axis is a major survival mechanism for pancreatic ductal adenocarcinoma under genotoxic stress. *Sci. Transl. Med.* **13**, 5445 (2021).

39. Dosset, M. et al. PD-1/PD-L1 pathway: an adaptive immune resistance mechanism to immunogenic chemotherapy in colorectal cancer. *Oncoimmunology* **7**, e1433981 (2018).

40. Bushnell, B. *BBMap: A Fast, Accurate, Splice-Aware Aligner.* Report No. LBNL-7065E (Lawrence Berkeley National Laboratory, 2014).

41. Beghini, F. et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *eLife* **10**, e65088 (2021).

42. Almeida, A. et al. A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* **39**, 105–114 (2021).

43. McMurdie, P. J. & Holmes, S. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).

44. Segata, N. et al. Segata-LEfSe-gb-2011. *Genome Biol.* **12**, R60 (2011).

45. Caporaso, J. G. et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6**, 1621–1624 (2012).

46. Callahan, B. J. et al. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

47. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

48. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

49. Jassal, B. et al. The reactome pathway knowledgebase. *Nucleic Acids Res.* **48**, D498–D503 (2020).

50. Kempski, J. et al. IL22BP mediates the antitumor effects of lymphotoxin against colorectal tumors in mice and humans. *Gastroenterology* **159**, 1417–1430 (2020).

Acknowledgements

We thank the Mouse Pathology Core Facility, the Bioinformatic Core Facility, the Cytometry und Cell Sorting Core Unit and the members of the Gnotobiotic Mice Facility at the University Medical Center Hamburg-Eppendorf for their technical assistance; the patients and their families, and the investigators and research staff at all study sites; C. Saygi for performing the differential expression analysis of the RNA-seq data; C. Bang for

technical assistance; C. Esser for *Ahr*$^{-/-}$ knockout bone marrow; S. Baldus, M. Mollenhauer, K. Tinaz and C. Vosen for *Mpo*$^{-/-}$ knockout bone marrow; M. Hamley for editing; and H. Pinnschmidt of the Institute of Medical Biometry and Epidemiology for statistical advice. This work is supported by a grant from the Brigitte und Dr. Konstanze Wegener-Stiftung to J.T.; German Cancer Aid fellowship grants to J.T., M. Schönlein and L.K. and grant 70114815 to J.T. and N.G.; ERC StG 715271 grant to N.G.; ERC COG 865466 grant to S.H.; and funding by the German Research Foundation (DFG) project 467261817 and SFB841 (via B. Fehse) to K.R. V.G.P. was supported by the Bundesministerium für Bildung und Forschung (eMed Consortia “Fibromap”) and the Novo Nordisk Foundation (Young Investigator Award; NNF21OC0066381). Furthermore, T.R.L., L.A. and T.S. are funded by VolkswagenStiftung’s initiative ‘Niedersächsisches Vorab’ (76251-99) and by the German Research Foundation under Germany’s Excellence Strategy (EXC 2155, project number 390874280). F.C. is supported by an iCARE-2 fellowship issued by AIRC and the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 800924. The work of J.T.S. is supported by the German Cancer Consortium (DKTK), by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through 405344257 (SI 1549/3–2) and SI1549/4–1, and by the German Federal Ministry of Education and Research (BMBF; 01KD2206A/SATURN3). M.T.-A. and J.T.S. are members of the German Cancer Consortium (DKTK). Figure 1a, Extended Data Fig. 2a,d and small icons throughout the figures were created with BioRender.com.

Funding

Open access funding provided by Universitätsklinikum Hamburg-Eppendorf (UKE).

Author information

Author notes

1. These authors jointly supervised this work: Samuel Huber, Nicola Gagliani

Authors and Affiliations

1. II. Department of Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

   Joseph Tintelnot, Martin Schönlein, Alexander Stein, Carsten Bokemeyer & Marianne Sinn

2. Mildred Scheel Cancer Career Center HaTriCS4, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

   Joseph Tintelnot, Martin Schönlein & Leonie Konczalla

3. Department of General, Visceral and Thoracic Surgery, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

   Yang Xu, Leonie Konczalla, Anastasios D. Giannou, Filippo Cortesi, Tobias S. Bröring, Faik G. Uzunoglu, Cenap Güngör, Jakob R. Izbicki & Nicola Gagliani

4. Research Group Microbial Immune Regulation, Helmholtz Centre for Infection Research, Braunschweig, Germany

   Till R. Lesker, Agata A. Bielecka, Lena Amend & Till Strowig

5. I. Department of Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

   Anastasios D. Giannou, Penelope Pelczar, Samuel Huber & Nicola Gagliani

6. Hamburg Center for Translational Immunology (HCTI), Hamburg, Germany

   Anastasios D. Giannou, Penelope Pelczar, Samuel Huber & Nicola Gagliani

7. III. Department of Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Dominik Kylies & Victor G. Puelles

8. Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

Victor G. Puelles

9. Institute of Clinical Chemistry and Laboratory Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Manuela Peschka & Thomas Renné

10. Newborn Screening and Metabolic Laboratory, Department of Pediatrics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Manuela Peschka

11. Research Department Cell and Gene Therapy, Department of Stem Cell Transplantation, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Kristoffer Riecken & Maximilian Jung

12. Bridge Institute of Experimental Tumor Therapy, West German Cancer Center, University Hospital Essen, University Duisburg-Essen, Essen, Germany

Marija Trajkovic-Arsic & Jens T. Siveke

13. Division of Solid Tumor Translational Oncology, German Cancer Consortium (DKTK Partner Site Essen) and German Cancer Research Center (DKFZ), Heidelberg, Germany

Marija Trajkovic-Arsic & Jens T. Siveke

14. Irish Centre for Vascular Biology, School of Pharmacy and Biomolecular Sciences, Royal College of Surgeons in Ireland, Dublin, Ireland

Thomas Renné

15. Center for Thrombosis and Hemostasis (CTH), Johannes Gutenberg University Medical Center, Mainz, Germany

Thomas Renné

16. Department of Internal Medicine III, Ludwig-Maximilians-University (LMU) Hospital, Munich, Germany

Danmei Zhang & Stefan Boeck

17. Hannover Medical School (MHH), Hannover, Germany

Till Strowig

18. Hematology–Oncology Practice Hamburg (HOPE), University Cancer Center Hamburg, Hamburg, Germany

Alexander Stein

19. Department of Radiation Oncology, Perlmutter Cancer Center, New York University Grossman School of Medicine, New York, NY, USA

Alec C. Kimmelman

Contributions

J.T. performed or supervised all experiments, performed most of the data analysis, wrote the manuscript and designed the project. Y.X. helped with experimental work and performed blinded tumour assessment. T.R.L. performed bioinformatic microbiota analysis. M. Schönlein helped with patient recruitment and editing of the manuscript. L.K., A.D.G. and P.P. helped to set up experimental mouse models and gnotobiotic mouse experiments. D.K. and V.G.P. acquired widefield microscopic images.

A.A.B. extracted and processed DNA for shotgun metagenomics. M.P. and T.R. set up and acquired metabolomic data. F.C. helped with flow cytometry analysis. K.R. and M.J. performed lentiviral shRNA knockdown. L.A. extracted and cultured bacterial strains from the stool of R patients. T.S.B. helped with experimental work. M.T.-A. and J.T.S. provided intellectual feedback and support. D.Z., S.B., F.G.U. and A.S. helped with patient recruitment. T.S. provided the shotgun metagenomic pipeline and intellectual feedback. C.G. provided cancer cell lines and advice. J.R.I. and C.B. provided intellectual feedback and support and edited the manuscript. M. Sinn provided clinical supervision and feedback, edited the manuscript and helped recruit patients. A.C.K. provided autophagy reporter cells and intellectual feedback and edited the manuscript. S.H. supervised the research and edited the paper. N.G. designed and supervised the project and wrote the manuscript.

Corresponding authors

Correspondence to Joseph Tintelnot or Nicola Gagliani.

Ethics declarations

Competing interests

D.Z. received travel support from AstraZeneca and AMGEN, honoraria from AstraZeneca and Roche. S.B. reports honoraria from AstraZeneca, BMS, Celgene, Incyte, Janssen Cilag, MSD and Servier for scientific presentations and paid consulting, as well as research support from Celgene. J.T.S. receives honoraria as consultant or for continuing medical education presentations from AstraZeneca, Bayer, Boehringer Ingelheim, Bristol-Myers Squibb, Immunocore, MSD Sharp Dohme, Novartis, Roche/Genentech, and Servier. His institution receives research funding from Abalos Therapeutics, Boehringer Ingelheim, Bristol-Myers Squibb, Celgene, Eisbach Bio, and Roche/Genentech; he holds ownership and serves on the Board of Directors of Pharma15. C.B. reports personal fees from Sanofi Aventis, Merck KgA, Bristol-Myers Squibb, Merck Sharp & Dohme, Lilly Imclone, Bayer Healthcare, GSO Contract Research, AOK Rheinland-Hamburg and Novartis. M. Sinn received honoraria from AstraZeneca,

Amgen, BMS, MSD, Incyte, Pierre Fabre, Pfizer Servier and Sanofi and support for clinical research (institutional) from Amgen, AstraZeneca, Bayer, BMS, Incyte, MSD, Pierre Fabre, Roche and Servier. N.G. reports financial support from Roche. All of these are outside the submitted work. A.C.K. has financial interests in Vescor Therapeutics, OncoRev and is an inventor on patents pertaining to KRAS-regulated metabolic pathways and redox control pathways in pancreatic cancer, targeting GOT1 as a therapeutic approach, targeting alanine transport, and the autophagic control of iron metabolism. A.C.K. is on the scientific advisory board of Rafael/Cornerstone Pharmaceuticals and has been a consultant for Deciphera and Abbvie. The remaining authors declare no competing interests.

Peer review

Peer review information

Nature thanks Filipe Cabreiro, Marco Cassatella and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Survival and microbiota analysis of R and NR patients with mPDAC from the Hamburg cohort.

a, 30 patients were recruited to the study. **b**,**c**, After exclusion of non-eligible patients as indicated, 23 patients could be analysed for PFS (**b**) and overall survival (OS) (**c**). PFS and OS are presented separately for responder (R, n = 11, blue) and non-responder (NR, n = 12, red) patients. **d**–**f**, After exclusion of one sample due to sample collection errors that led to sequencing failure, the microbiota of 22 patients was analysed. **d**,

Microbiota LEfSe analysis showing linear discriminant analysis (LDA) score of bacterial taxa enriched in R or NR patients, respectively. **e**, Genus heat map with Ward clustering of patients. **f**, Microbiota diversity compared using Shannon index. Boxes indicate 25 to 75% of values (**f**) and error bars indicate median, significant p-values are indicated and were determined by Gehan–Breslow–Wilcoxon test (**b**,**c**) or two-tailed Wilcoxon test (**f**).

Source data

Extended Data Fig. 2 R microbiota induces a response to FIRINOX treatment.

a, Microbiota of five R and NR patients with mPDAC was collected and transferred to gnotobiotic mice. Nine patients were treated with FIRINOX and one with GnP. PFS (**b**) and OS (**c**) of patients used for microbiota transfer experiments is depicted in Kaplan-Meier estimators. In four individual experiments with varying donors, one group of mice was colonized with R microbiota and one group was colonized with NR microbiota. After stool transfer, gnotobiotic mice (**d**) received orthotopic injections of KPC tumour cells and were either left untreated (n = 22 (NR) or 20 (R)) or treated once with FIRINOX (n = 23 (NR) or 20 (R)) at day eleven of the experiment. **e**, Tumour weight of orthotopic tumours is depicted for the respective experimental group at day 20 of the experiment. **f**, Randomly selected tumours from varying experiments of e were analysed for intratumoral bacteria using 16S rRNA sequencing (n = 12). Table shows indicated clean reads for different tumours (T1–12), H_2O or positive control. **g**, 3-IAA was analysed in the serum of n = 3 to 5 randomly selected mice per experiment with different donors each. **h**, Metagenomic sequencing data was analysed for the abundance of a panel of 3-IAA-producing bacteria. Increased abundance of *B. fragilis* and *B. thetaiotaomicron* was detected in stool samples from NR or R patients during and before chemotherapy treatment, which were used for experiments in gnotobiotic mice (5 donors per group and 1–3 samples per patient, n = 12 per group). **i**, 3-IAA was quantified in supernatants of bacteria lacking (*Prevotella copri*) and having the ability (*B. thetaiotaomicron* and *B. fragilis*) to produce 3-IAA based on genomic predictions. The characterized strains included *P. copri* (DSM18205), *B. fragilis* (BF DSM2151, Bf0901, Bf0902) and *B.*

thetaiotaomicron (DSM, Bt0901, Bt0902). Strains were all cultured in BHI containing 1% tryptophan, the strain BFi2 was also cultured in BHI alone (n = 3 biological replicates). **j**, SPF mice were orthotopically injected with KPC cells. Mice were either supplemented with tryptophan-free, standard diet (2.3 g/kg TRP) or tryptophan-high diet (12 g/kg TRP) starting 7 days after tumour cell injection for a total of 14 days (n = 3). **k**, As in **j**, except dietary intervention was applied for a total of 4 days before standard diet was reintroduced. Tumour weight was assessed 9 days after dietary intervention (n = 3 or 4). **l**, NR microbiota-colonized mice (red) were fed either standard diet (SD) or tryptophan-high diet for 4 days (TRP high). Blood was drawn at day 4 of intervention and 3-IAA was measured in the serum using 3-IAA CLIA (n = 4). **m**, Tumour sizes of mice from **l** at day nine after treatment with FIRINOX are shown (n = 3 or 4). Each symbol represents one mouse, human or *in vitro* replicate. One experiment was performed (**f**,**i**–**m**) or four (**e**,**g**) independent experiments were pooled. Error bars indicate SEM, whiskers indicate 10% and 90% (**h**), significant p-values are indicated and were determined by Gehan-Breslow-Wilcoxon test (**b**,**c**), one-way ANOVA followed by Tukey's (**e**,**m**) or Dunnett's (**i**–**k**) post-hoc test, two-way ANOVA (**g**), two-tailed Wilcoxon paired-ranks test (**h**) or two-tailed t-test (**l**).

Source data

Extended Data Fig. 3 3-IAA induces a response to FIRINOX in mouse models of PDAC.

a, 3-IAA serum concentration of SPF mice gavaged with 500 mg/kg 3-IAA was measured using CLIA at 2, 6 and 24 h after 3-IAA application (n = 3 mice). Unrelated serum concentration of R-microbiota-colonized gnotobiotic mice is shown as a control. **b**, SPF mice were orthotopically injected with KPC cells and received two treatments of FIRINOX or FIRINOX + 3-IAA (n = 5). Tumour weight at day 20 of the experiment is depicted in the statistic. **c**, NR-colonized gnotobiotic mice were injected with KPC tumour cells orthotopically. 9 days later, mice were substituted +/− 3-IAA for five days (d9–13) and treated +/− FIRINOX as depicted in the experimental scheme (n = 3 or 4). Tumour weight is shown at day 20 of the experiment. **d**, SPF mice were orthotopically injected with KPC tumour cells. At day 9 after

tumour cell injection, mice were treated with either 500 mg/kg 3-IAA, 3-IPA or 250 mg/kg GCA and 250 mg/kg DCA for five consecutive days (n = 4). FIRINOX treatment was applied at day 11. Tumour weight is shown at day 20 of the experiment. **e**, As in **d**, except mice were treated with FIRINOX, FIRINOX + 3-IAA (500 mg/kg) or FIRINOX + hippuric acid (500 mg/kg) (n = 4 or 5). Each symbol represents one mouse. One experiment (**e**) or one out of two independent experiments is shown (**a**–**d**). Error bars indicate SEM, significant p-values are indicated and were determined by Kruskal–Wallis test followed by Dunn's post-hoc test (**a**,**c**,**e**), or one-way ANOVA followed by Tukey's (**b**) or Dunnett's (**d**) post-hoc test.

Source data

Extended Data Fig. 4 Tumour-infiltrating immune cells of KPC tumours from gnotobiotic mice.

a, Flow cytometric gating strategy to classify immune subsets (left panel) or intracellular cytokines and coinhibitory receptors (right panel). Immune subsets were determined using the following marker combinations on live $CD45^{+}$ immune cells: $CD8^{+}$ T cells ($CD11b^{-}CD3^{+}CD8^{+}$); $CD4^{+}$ T cells ($CD11b^{-}CD3^{+}CD4^{+}$); neutrophils ($CD11b^{+}Ly6G^{+}$); and macrophages ($CD11b^{+}Ly6G^{-}F4/80^{+}$). **b**, Immune subsets or cytokine profiles and PD-1 expression of immune cells from orthotopic tumours obtained from gnotobiotic mice colonized with R or NR microbiota was analysed at day 20 of the experiment. Each symbol represents one mouse. Three to five independent experiments were pooled (n = 11 to 23 in total). Error bars indicate SEM, significant p-values are indicated and were determined by two-tailed t-test.

Source data

Extended Data Fig. 5 $CD4^{\pm}$ and $CD8^{\pm}$ T cells are dispensable for the efficacy of 3-IAA and FIRINOX.

R-microbiota-colonized gnotobiotic mice were injected with KPC tumour cells orthotopically. Before treatment with FIRINOX, mice were injected with either isotype control antibody or CD8-depleting antibody every third

day, as indicated. The tumour weight (**a**) and the depletion efficacy of intratumoral $CD8^+$ T cells (**b**) as relative to total immune cells ($CD45^+$), determined by flow cytometry, is shown at day 20 of the experiment (n = 3). **c**, As in **a**, except that $CD4^+$ and $CD8^+$ T cells were depleted simultaneously (n = 3 or 4). **d**, R-microbiota-colonized gnotobiotic mice with orthotopic KPC tumours were treated with FIRINOX, FIRINOX + tryptophan-high diet (d8–12) + isotype control antibody or FIRINOX + tryptophan-high diet (d8–12) + CD4/CD8-depleting antibody (n = 4 or 5). Tumour weight at day 20 of the experiment is depicted. **e**, As in **c**, except that SPF mice were treated with 3-IAA +/− FIRINOX and CD4/CD8-depleting or isotype control antibody (n = 4). Tumour weight is depicted at day 20 of the experiment. Each symbol represents one mouse. One experiment each was performed. Error bars indicate SEM, significant p-values are indicated and were determined by two-tailed t-test (**a**,**b**), one-way ANOVA followed by Dunnett's (**c**,**d**) post-hoc test or Kruskal–Wallis test followed by Dunn's post-hoc test (**e**).

Source data

Extended Data Fig. 6 3-IAA and chemotherapy induce cell death in neutrophils.

a, FACS-sorted T cells or neutrophils were cultured for 24 h +/− 1,000 µM 3-IAA and +/− FIRINOX (3.2 µM Oxaliplatin, 5.6 µM Irinotecan and 19.2 µM 5-FU). Viability was assessed via flow cytometry (n = 3, biological replicates). **b**, MPO activity was determined in 50,000 FACS-sorted bone-marrow-derived neutrophils, neutrophil precursors (lineage-negative, $CD115^-$, $Ly6B^+$, $Ly6G^{int\text{-}low}$) from the bone marrow or PDAC-infiltrating neutrophils using a fluorometric MPO activity assay kit (n = 4, biological replicates). **c**, FACS-sorted bone-marrow-derived neutrophils were cultured in the presence of increasing dosages of 3-IAA +/− MPO +/− FIRINOX (3.2 µM Oxaliplatin, 5.6 µM Irinotecan and 19.2 µM 5-FU; n = 3, biological replicates). Cell numbers were quantified using flow cytometry after 48 h of culture. **d**, FACS-sorted neutrophil precursors were cultured in the presence of increasing dosages of 3-IAA +/− FIRINOX (3.2 µM Oxaliplatin, 5.6 µM Irinotecan and 19.2 µM 5-FU; n = 3, biological replicates). Cell numbers were quantified using flow cytometry after 48 h of culture. **e**, As in **a**, except

only neutrophils were cultured in the presence of +/− Oxaliplatin (8 µM Oxaliplatin) and 1,000 µM 3-IAA, 1,000 µM 3-IPA or DMSO at similar concentrations (n = 3 to 6, biological replicates). **f**, As in **c**, except cell death of neutrophils was defined by flow cytometry using Annexin V and PI staining. Frequencies are shown relative to total neutrophils (n = 3, biological replicates). **g**, As in **c**, except NET formation was defined by SYTOX DNA staining after three hours of treatment (n = 3, biological replicates). **h**, Degranulation of 1×10^6 FACS-sorted bone-marrow-derived neutrophils was measured after 30 min of stimulation with FIRINOX (3.2 µM Oxaliplatin, 5.6 µM Irinotecan and 19.2 µM 5-FU), 1 mM 3-IAA + FIRINOX or FIRINOX + 1 µM fMLP using an MPO activity assay kit (*n* = 3, biological replicates). **i**, KPC tumour cells were orthotopically injected into SPF mice and mice were treated +/− 500 mg/kg 3-IAA for 5 consecutive days and FIRINOX at day 11 (n = 5 each). Immune cells from the spleen and tumour were analysed by flow cytometry at day 3 after FIRINOX treatment. Neutrophil counts in the spleen or tumour are shown. **j**, As in **i**, except only 3-IAA, but no FIRINOX was applied (n = 5 each). Each symbol represents one mouse or one in vitro replicate. One experiment (**b**) or one out of three (**c**,**d**) or two (**a**,**e**–**j**) independent experiments are shown. Error bars indicate SEM, significant p-values are indicated and were determined by one-way ANOVA followed by Dunnett's (**a**–**h**) post-hoc test or two-tailed t-test (**i**,**j**).

Source data

Extended Data Fig. 7 3-IAA and MPO increase ROS and decrease cell viability.

a, Concentrations of 3-IAA measured using LC–MS/MS (left) or MOI (right) of tumours isolated from SPF mice treated with FIRINOX +/− 3-IAA five hours after treatment (n = 3 biological replicates). **b**, Concentrations of 3-IAA (left) or MOI (right) were measured in tumours isolated from WT or *Mpo*$^{-/-}$ bone-marrow-reconstituted mice treated with FIRINOX + 3-IAA using LC–MS/MS five hours after treatment (n = 3 or 4 biological replicates). **c**, SPF mice received KPC control or *Ahr* KD cells orthotopically and were treated with FIRINOX (d11) or FIRINOX + 3-IAA (d9–13) (n = 4 or 5). Tumour weight was assessed at day 20 of the

experiment. **d**, KPC tumour cells were cultured in the presence of increasing dosages of 3-IAA +/− Oxaliplatin (8 μM Oxaliplatin; n = 2-3 biological replicates). ROS expression was determined via flow cytometry using CellROX dye. Hy19636 cells were cultured with increasing dosages of 3-IAA +/− 5 × 10^4 neutrophils (**e**) or MPO 200mU/ml (**f**) and +/− FIRINOX (1.6 μM Oxaliplatin, 2.8 μM Irinotecan and 9.6 μM 5-FU; n = 2-3 biological replicates). ROS expression was assessed via flow cytometry. Only p-values for relevant groups are shown for clarity. MIA PaCa-2 (**g**), T3M-4 (**h**) or Hy19636 (**i**) cells were cultured with increasing dosages of 3-IAA +/− FIRINOX (3.2 μM Oxaliplatin, 5.6 μM Irinotecan and 19.2 μM 5-FU) for 6 h (n = 3 biological replicates). Viability was assessed by MTS/MTT assay and displayed as relative to untreated cells after 48 h. KPC (**j**) or MIA PaCa-2 (**k**) cells were cultured in the presence of increasing dosages of 3-IAA +/− MPO 400 mU/ml +/− FIRINOX (3.2 μM Oxaliplatin, 5.6 μM Irinotecan and 19.2 μM 5-FU; n = 3 biological replicates). Cell numbers were quantified using flow cytometry after 48 h of culture. Hy19636 (**l**) or MIA PaCa-2 (**m**) cells were cultured with increasing dosages of MOI (100, 200 or 400 μM) +/− FIRINOX (1.6 μM Oxaliplatin, 2.8 μM Irinotecan and 9.6 μM 5-FU; n = 5 to 12 biological replicates). Viability was assessed using the MTS assay and displayed as relative to untreated cells after 24 h. **n**, KPC tumour cells were orthotopically injected into SPF mice and mice were left untreated or treated with either FIRINOX (d11), 3-IAA (d9–13) or 3-IAA and FIRINOX (n = 5 each). Statistics and representative pictures show orthotopic tumours three days after indicated treatment with staining for nitrotyrosine (ROS). **o**, SPF mice were injected with orthotopic tumours and were treated with FIRINOX or FIRINOX + 3-IAA (n = 5 each). After five hours of treatment, ROS accumulation was analysed using flow cytometry. ROS levels of cancer cells ($Epcam^+$), T cells ($CD45^+CD3^+$) or myeloid cells ($CD45^+CD11b^+$) are shown. **p**, Irradiated and WT (n = 5) or *Mpo*$^{-/-}$ (n = 3) bone-marrow-reconstituted mice received KPC cells orthotopically and were treated with FIRINOX or FIRINOX + 3-IAA as in **o**. ROS levels were assessed by flow cytometry as in **o**. One (**a**,**b**,**n**) or one out of three (**d**,**g**–**i**) or two (**c**,**e**,**f**,**j**–**m**,**o**,**p**) independent experiments are shown. Error bars indicate SEM or median (**n**), significant p-values are indicated and were determined by two-tailed t-test (**a**–**c**,**o**,**p**), one-way ANOVA followed by Dunnett's (**d**,**g**–**n**) or Tukey's (**e**,**f**) post-hoc test.

Source data

Extended Data Fig. 8 3-IAA and FIRINOX treatment decreases autophagy in vivo.

a, Normalized expression of ROS-producing or degrading enzymes found in mRNA sequencing data of tumours obtained from mice treated with FIRINOX compared to tumours treated with FIRINOX + 3-IAA (n = 3 biological replicates each; p values: GPX3 (0.00012), GPX7 ($1.2*10^{-8}$), NOX4 (0.023)). **b**, KPC tumour cells were cultured in the presence of FIRINOX (3.2 µM Oxaliplatin, 5.6 µM Irinotecan and 19.2 µM 5-FU) +/− 10 µM of 3-IAA and 400 mU/ml MPO (n = 4 biological replicates). RNA expression of GPX 3 and 7 was measured using qPCR after 24 h of treatment. Expression is depicted as relative to a housekeeping gene. **c**, Scramble control-transfected cells, *Gpx3* or *Gpx7* knockdown cells were treated with +/− FIRINOX (3.2 µM Oxaliplatin, 5.6 µM Irinotecan and 19.2 µM 5-FU) +/− 10 µM of 3-IAA and MPO and ROS was assessed using flow cytometry (n = 3 or 4 biological replicates). **d**, as in **c**, except cell numbers were calculated using flow cytometry (n = 4 biological replicates). **e**, SPF mice were injected with orthotopic tumours using scramble control or GPX3 KD cells and were treated with +/− FIRINOX at day eleven of the experiment (n = 5 or 6). Tumour size was measured at day eight after FIRINOX treatment. **f**, GSEA enrichment plot depicting positive enrichment of reactome pathway autophagy in total tumours obtained from mice treated with FIRINOX compared to tumours treated with FIRINOX and 3-IAA (n = 3 biological replicates). **g**, KPC tumour cells were orthotopically injected into SPF mice and mice were treated with either 3-IAA and FIRINOX or FIRINOX alone (n = 3 each). One day after FIRINOX treatment, intratumoral proteins were analysed and compared. Proteins downregulated in tumours from 3-IAA and FIRINOX-treated mice were analysed for enriched KEGG pathways. The top 10 pathways are depicted and the full list of proteins and up- or downregulated pathways is provided in SI tables 1 to 3. **h**, KPC tumour cells were orthotopically injected into SPF mice. Mice were treated +/− FIRINOX (d11), +/− 3-IAA (d9–13) and p62/SQSTM1 staining was analysed at day three after treatment. Statistics show the number of p62/SQSTM1^{high} cells per field (n = 5 each). Representative pictures are shown. The scale bar represents 100 µm. **i**,**j**, KPC tumour cells

were orthotopically injected into gnotobiotic mice colonized with R or NR microbiota (n = 3 or 4). Mice were treated +/− FIRINOX and LC3-l/ll (**i**) or p62/SQSTM1 (**j**) staining was analysed at day 20 of the experiment. Statistics show the number of LC3-l/ll^{high} or p62/SQSTM1^{high} cells per field (n = 4). Tumours were pooled from three individual experiments with different donors each. Representative pictures are shown. The scale bar represents 100 µm. **k**, KPC tumour cells were orthotopically injected into SPF mice and treated and analysed as in **h**, except that staining for cleaved caspase-3 (CC3, apoptosis) was applied. The scale bar represents 50 µm. **l**, SPF mice were orthotopically injected with GFP-LC3B-RFP reporter cells and treated with FIRINOX + 3-IAA, trehalose or FIRINOX + 3-IAA + trehalose (n = 5 each). Tumours were analysed for GFP/RFP ratios via bright field imaging. The GFP/RFP ratio is shown for each respective group. **m**, SPF mice were injected with (mST-ATG4B) or mSt control cells. Mice received either doxycycline treatment via diet for seven days (d5 to d12) or standard food and were treated +/− FIRINOX (n = 4 or 5 each). Tumour size at day 18 of the experiment is depicted. Representative pictures show tumours at the end of the experiment. Scale bar shows one centimetre. One (**a**,**f**–**k**) or one out of two independent experiments (**b**–**e**,**l**) or three independent experiments with different lengths of doxycycline treatment (**m**) are shown. Each symbol represents one mouse or one *in vitro* replicate. Bars indicate SEM (**a**–**e**,**m**) or median (**h**–**l**), significant p-values are indicated and were determined as indicated in the methods section (**a**,**f**,**g**), by two-tailed Mann–Whitney test (**b**) one-way ANOVA followed by Tukey's (**c**–**e**,**i**,**j**,**m**) or Dunnett's (**l**) post-hoc test or by Kruskal–Wallis test followed by Dunn's (**h**,**k**) post-hoc test.

Source data

Extended Data Fig. 9 A 3-IAA and chemotherapy combination is effective in the treatment of colorectal and lung cancer cell lines.

a,**b**, The MC38 colorectal cancer cell line (n = 5 or 6; **a**) or LLC lung cancer cell line (n = 4 or 5; **b**) was subcutaneously injected and mice were treated +/− FIRINOX and +/− 3-IAA (500 mg/kg) orally. Tumours were measured every other day and tumour weight was scaled to day 17 of experiment. **c**,

Orthotopic KPC tumours were established and SPF mice were treated with GNP once +/− 3-IAA (500 mg/kg; n = 4 each). Tumour weight was assessed seven days after chemotherapy (d18). **d**, Statistics show the event of neutropenia (<1,5*10^9/l) during the first 6 cycles of chemotherapy in patients from the Hamburg cohort responding or not responding to the chemotherapy (n = 11 R and 12 NR). **e**, DCA concentration was measured after two to three chemotherapy cycles in the serum of patients from the Hamburg cohort and was correlated with PFS (n = 20) or overall survival (n = 18). Each symbol represents one patient. **f**, Univariate Cox proportional hazard models were used to determine the effects of indicated variables on PFS in the Hamburg cohort. One experiment (**b**) or one out of two individual experiments (**a**) is shown. Error bars indicate SEM, significant p-values are indicated and were determined by one-way ANOVA followed by Dunnett's post-hoc test (**a**, left statistic), Kruskal–Wallis test followed by Dunn's (**b**, left statistic) post-hoc test, mixed-effects analysis followed by Dunnett's post-hoc test (**a** and **b**, right statistic), two-tailed t-test (**c**), two-sided chi-square test (**d**), simple linear regression and Pearson's *r* (**e**) or Cox regression (**f**).

Source data

Extended Data Table 1 Patient data for the Hamburg (observation) and Munich (validation) cohorts

Supplementary information

Reporting Summary

Supplementary Table 1

Protein screen comparing tumours isolated from 3-IAA and FIRINOX-treated mice to those isolated from FIRINOX-only-treated mice (control). Protein screen was performed and analysed as described in the Methods.

Supplementary Table 2

KEGG pathways of proteins downregulated in tumours isolated from 3-IAA and FIRINOX-treated mice. Proteins with a difference of < −0.35 compared to FIRINOX-only-treated tumours (Table 1) were selected and analysed as described in the Methods.

Supplementary Table 3

KEGG pathways of proteins upregulated in tumours isolated from 3-IAA and FIRINOX-treated mice. Proteins with a difference of > +0.35 compared to FIRINOX-only-treated tumours (Table 1) were selected and analysed as described in the Methods.

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Source Data Extended Data Fig. 1

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 3

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 6

Source Data Extended Data Fig. 7

[Source Data Extended Data Fig. 8]

[Source Data Extended Data Fig. 9]

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Tintelnot, J., Xu, Y., Lesker, T.R. *et al.* Microbiota-derived 3-IAA influences chemotherapy efficacy in pancreatic cancer. *Nature* **615**, 168–174 (2023). https://doi.org/10.1038/s41586-023-05728-y

- Received: 15 February 2022
- Accepted: 12 January 2023
- Published: 22 February 2023
- Issue Date: 02 March 2023

- DOI: https://doi.org/10.1038/s41586-023-05728-y

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

This article is cited by

- **A gut reaction can tune tumour fate during chemotherapy**
  - Le Li
  - Florencia McAllister

  Nature (2023)

A gut reaction can tune tumour fate during chemotherapy

- Le Li
- Florencia McAllister

Nature News & Views 22 Feb 2023

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05728-y

Amendments & Corrections

- **Author Correction: Common and rare variant associations with clonal haematopoiesis phenotypes** [17 Feb 2023]
-

- Author Correction
- Published: 17 February 2023

Author Correction: Common and rare variant associations with clonal haematopoiesis phenotypes

- Michael D. Kessler,
- Amy Damask,
- Sean O'Keeffe,
- Nilanjana Banerjee,
- Dadong Li,
- Kyoko Watanabe,
- Anthony Marketta,
- Michael Van Meter,
- Stefan Semrau,
- Julie Horowitz,
- Jing Tang,
- Jack A. Kosmicki,
- Veera M. Rajagopal,
- Yuxin Zou,
- Yariv Houvras,
- Arkopravo Ghosh,
- Christopher Gillies,
- Joelle Mbatchou,
- Ryan R. White,
- Niek Verweij,
- Jonas Bovijn,
- Neelroop N. Parikshak,
- Michelle G. LeBlanc,
- Marcus Jones,
- Regeneron Genetics Center,

- GHS-RGC DiscovEHR Collaboration,
- David J. Glass,
- Luca A. Lotta,
- Michael N. Cantor,
- Gurinder S. Atwal,
- Adam E. Locke,
- Manuel A. R. Ferreira,
- Raquel Deering,
- Charles Paulding,
- Alan R. Shuldiner,
- Gavin Thurston,
- Adolfo A. Ferrando,
- Will Salerno,
- Jeffrey G. Reid,
- John D. Overton,
- Jonathan Marchini,
- Hyun M. Kang,
- Aris Baras,
- Gonçalo R. Abecasis &
- …
- Eric Jorgenson

Show authors

Nature **volume 615**, page E3 (2023)

- 418 Accesses
- 10 Altmetric
- Metrics details

The Original Article was published on 30 November 2022

Correction to: *Nature* https://doi.org/10.1038/s41586-022-05448-9
Published online 30 November 2022

In the version of this article originally published, Extended Data Fig. 7d was an inadvertent duplication of Extended Data Fig. 7a and has now been corrected in the HTML and PDF versions of the article.

Author information

Author notes

1. A full list of members and their affiliations appears in the Supplementary Information

Authors and Affiliations

1. Regeneron Genetics Center, Tarrytown, NY, USA

   Michael D. Kessler, Amy Damask, Sean O'Keeffe, Nilanjana Banerjee, Dadong Li, Kyoko Watanabe, Anthony Marketta, Julie Horowitz, Jing Tang, Jack A. Kosmicki, Veera M. Rajagopal, Yuxin Zou, Arkopravo Ghosh, Christopher Gillies, Joelle Mbatchou, Niek Verweij, Jonas Bovijn, Neelroop N. Parikshak, Michelle G. LeBlanc, Marcus Jones, Luca A. Lotta, Michael N. Cantor, Adam E. Locke, Manuel A. R. Ferreira, Charles Paulding, Alan R. Shuldiner, Adolfo A. Ferrando, Will Salerno, Jeffrey G. Reid, John D. Overton, Jonathan Marchini, Hyun M. Kang, Aris Baras, Gonçalo R. Abecasis & Eric Jorgenson

2. Regeneron Pharmaceuticals, Tarrytown, NY, USA

   Michael Van Meter, Stefan Semrau, Yariv Houvras, Ryan R. White, David J. Glass, Gurinder S. Atwal, Raquel Deering & Gavin Thurston

Consortia

Regeneron Genetics Center

- Michael D. Kessler

- , Amy Damask
- , Sean O’Keeffe
- , Nilanjana Banerjee
- , Dadong Li
- , Kyoko Watanabe
- , Anthony Marketta
- , Julie Horowitz
- , Jing Tang
- , Jack A. Kosmicki
- , Veera M. Rajagopal
- , Yuxin Zou
- , Yariv Houvras
- , Arkopravo Ghosh
- , Christopher Gillies
- , Joelle Mbatchou
- , Niek Verweij
- , Jonas Bovijn
- , Neelroop N. Parikshak
- , Michelle G. LeBlanc
- , Marcus Jones
- , Luca A. Lotta
- , Michael N. Cantor
- , Adam E. Locke
- , Manuel A. R. Ferreira
- , Charles Paulding
- , Alan R. Shuldiner
- , Adolfo A. Ferrando
- , Will Salerno
- , Jeffrey G. Reid
- , John D. Overton
- , Jonathan Marchini
- , Hyun M. Kang
- , Aris Baras
- , Gonçalo R. Abecasis
- & Eric Jorgenson

GHS-RGC DiscovEHR Collaboration

- Marcus Jones

Corresponding author

Correspondence to Eric Jorgenson.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Kessler, M.D., Damask, A., O’Keeffe, S. *et al.* Author Correction: Common and rare variant associations with clonal haematopoiesis phenotypes. *Nature* **615**, E3 (2023). https://doi.org/10.1038/s41586-023-05803-4

- Published: 17 February 2023
- Issue Date: 02 March 2023
- DOI: https://doi.org/10.1038/s41586-023-05803-4

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05803-4